DYNAMICAL GENERALIZATIONS OF THE LAGRANGE SPECTRUM

Sébastien Ferenczi

DYNAMICAL GENERALIZATIONS OF THE LAGRANGE SPECTRUM

SÉBASTIEN FERENCZI

Abstract. We compute two invariants of topological conjugacy, the upper and lower limits of the inverse of Boshernitzan's ne_n, where e_n is the smallest measure of a cylinder of length n, for three families of symbolic systems, the natural codings of rotations and three-interval exchanges and the Arnoux-Rauzy systems. The sets of values of these invariants for a given family of systems generalize the Lagrange spectrum, which is what we get for the family of rotations with the upper limit of $\frac{1}{ne_n}$.

The *Lagrange spectrum* is the set of finite values of $L(\alpha)$ for all irrational numbers α, where $L(\alpha)$ is the largest constant c such that $|\alpha - \frac{p}{q}| \leq \frac{1}{cq^2}$ for infinitely many integers p and q (a variant is known as the Markov spectrum, see Section 1.3 below). It was recently remarked that this arithmetic definition can be replaced by a dynamical definition involving the irrational rotations of angle α, through their natural coding by the partition $\{[0, 1-\alpha[, [1-\alpha, 1[\}$. Namely, as we prove in Theorem 2.4 below which was never written before, $L(\alpha)$ is also the upper limit of the inverse of the so-called *Boshernitzan's* ne_n, where e_n is the smallest (Lebesgue) measure of the nonempty cylinders of length n.

Thus, for any symbolic dynamical system, it is interesting to compute two new invariants of topological conjugacy, $\limsup_{n\to+\infty} \frac{1}{ne_n}$ and $\liminf_{n\to+\infty} \frac{1}{ne_n}$. Moreover, for a given family of systems, the set of all values of these invariants can be called the *upper*, resp. *lower BL* (for Boshernitzan and Lagrange) *spectrum*. In this paper, we compute these spectra for three families of systems: the irrational rotations (seen as two-interval exchanges), the three-interval exchanges, both coded by the natural partition of the interval generated by the discontinuities, and the Arnoux-Rauzy systems. In each of these cases, we use an induction (or renormalization) process, which is respectively a variant of the Euclid algoritm, the self-dual induction of [20], and the natural one defined in [3]. A multiplicative form of the process yields explicit formulas for our invariants, and these formulas are then exploited in each case by using the underlying algorithm of approximation of real numbers by rationals, which is respectively the classical continued fraction expansion, an extension of a semi-regular continued fraction expansion, and the algorithm which motivated the study of Arnoux-Rauzy systems.

What we get in the end is a first partial description of the five new sets we introduced beside the classical Lagrange spectrum. For rotations, the lower BL spectrum is a compact set starting with 1 and an interval (at least) as far as $1,03...$, ending at $1,38...$, with gaps, above an accumulation point at $1,23....$ For three-interval exchanges, the upper BL spectrum looks, perhaps deceptively, like two times the Lagrange spectrum, starting at $2\sqrt{5}$ with gaps and an accumulation point at 6, and ending with an interval (at least) from $14,8...$ to infinity; the lower BL spectrum is fully determined and is none other than the interval $[2, +\infty]$. For

Date: October 18, 2011.
1991 *Mathematics Subject Classification.* 37B10, 11J06, 68R15.

Arnoux-Rauzy systems, we deal with cubic numbers and our knowledge is only embryonic: the upper BL spectrum starts at $8,44...$, with gaps, and ends at infinity, the lower BL spectrum starts at 2, ends at infinity, and contains at least all the integers as accumulation points.

As a consequence, we get new uniquely ergodic systems for which ne_n does not tend to zero when n tends to infinity, showing that Boshernitzan's criterion (see Section 1.4) is not a necessary condition; their existence in the family of three-interval exchanges was known but a proof was never written, while the examples in the family of Arnoux-Rauzy systems are new, and surprising as these systems are often thought to behave like rotations, see the discussion at the end of Section 4. An interesting open problem would be to compute the BL spectra of the family of all uniquely ergodic symbolic systems, see Section 5.

1. Preliminaries

1.1. **Languages.**

Definition 1.1. *We look at finite* words *on a finite alphabet $\mathcal{A}$. A word with r letters, $w_1...w_r$, is of* length *r. The* concatenation *of two words w and w' is denoted by ww'. The* empty word *is the unique word of length zero.*
A word $w = w_1...w_r$ occurs at place *i in a word $v = v_1...v_s$ or an infinite sequence $v = v_1v_2...$ if $w_1 = v_i$, ..., $w_r = v_{i+r-1}$. We say that w is a* factor *of v. The empty word is a factor of any v.* Prefixes *and* suffixes *are defined in the usual way.*
A language *L is a set of words such if w is in L, all its factors are in L, aw is in L for at least one letter a of $\mathcal{A}$, and wb is in L for at least one letter b of $\mathcal{A}$.*
A language L is uniformly recurrent *if for each w in L there exists n such that w occurs in each word of length n of L.*

A language L is now fixed.

Definition 1.2. *A word w is* right special, *resp.* left special *if there exist at least two different letters x such that xw, resp. wx, is in L. If w is both right special and left special, w is* bispecial.
The complexity *of L is the function p_L which to each positive integer n associates the number of different words of length n in L.*
The Rauzy graph *of length n of L is the graph whose vertices are the words of length n in L, with an edge $w \to w'$ if there exists a word v of length $n-1$ such that $w = av$, $w' = vb$, and $avb \in L$.*

1.2. **Symbolic dynamics.**

Definition 1.3. *The* symbolic dynamical system *associated to a language L is the one-sided shift $S(x_0x_1x_2...) = x_1x_2...$ on the subset X_L of $\mathcal{A}^{\mathbb{N}}$ made with the infinite sequences such that for every $r < s$, $x_r...x_s$ is in L.*
For a word $w = w_1...w_r$ in L, the cylinder *$[w]$ is the set $\{x \in X_L; x_0 = w_1, ...x_{r-1} = w_r\}$.*
(X_L, S) is minimal *if L is uniformly recurrent.*
(X_L, S) is uniquely ergodic *if there is one S-invariant probability measure μ; then the* frequency *of the word w is the measure $\mu[w]$.*

Starting from a (in general, geometric in origin) topological dynamical system (X, T), we can get a symbolic dynamical system:

Definition 1.4. *For a transformation T defined on a set X, partitioned into X_1, ... X_r, and a point x in X, its* trajectory *is the infinite sequence $(x_n)_{n\in\mathbb{N}}$ defined by $x_n = i$ if $T^n x$ falls into X_i, $1 \leq i \leq r$.*
The language $L(T)$ is the set of all finite factors of its trajectories.
The coding *of (X,T) by the partition $\{X_1, ...X_r\}$ is the symbolic dynamical system $(X_{L(T)}, S)$.*

We shall often assimilate a dynamical system (X,T) with its coding by a partition. Note that the word $w_1...w_n$ has length n; in most cases, the cylinder $[w_1...w_n]$ is assimilated to an interval J, on which μ is the Lebesgue measure; the (geometrical) length of J is also the Lebesgue measure of $[w_1...w_n]$, and thus the frequency of the word $w_1...w_n$, which should not be mistaken with its (symbolic) length.

If the transformation T is minimal (i.e. every orbit is dense), all its trajectories have the same finite factors, and the language $L(T)$ is uniformly recurrent; the special words depend on the language and not on the individual trajectories; thus they are defined by any trajectory of T. If there is no periodic orbit, every word w is a factor of a bispecial word; hence the bispecial words determine the finite factors of the trajectories, and thus the symbolic dynamical system $(X_{L(T)}, S)$.

1.3. **Continued fractions.** In Section 3 below, we shall use the generalized, or *semi-regular continued fraction expansion*

$$\cfrac{1}{a_1 + \varepsilon_1 \cfrac{1}{a_2 + \varepsilon_2 \cfrac{1}{a_3 + \cdots}}}$$

for a_i positive integers, $\varepsilon_i = -$ or $+$, $\varepsilon_i = +$ if $a_i = 1$. We denote the above expression by

$$[0, a_1 * \varepsilon_1, a_2 * \varepsilon_2, ...].$$

We write a_n with no sign if the expansion stops at a_n. In Section 2 we shall deal only with classic (Euclid) continued fraction expansions, and write only a_i to denote $a_i * +$. The periodic sequence $a_1 * \varepsilon_1, ..., a_n * \varepsilon_n, a_1 * \varepsilon_1, ..., a_n * \varepsilon_n, , ...$ is denoted by $(a_1 * \varepsilon_1, ..., a_n * \varepsilon_n)^\omega$.

It is shown in [15] that $[0, a_1 * \varepsilon_1, ...]$. has bounded partial quotients (for the Euclid algorithm) if and only if the a_i are bounded and the number of consecutive $2 * -$ is bounded.

For a full study of the Lagrange spectrum, we refer the reader to the monograph [14]; the following definition is equivalent to the one given in the introduction of the present paper.

Definition 1.5. *The* Lagrange spectrum *is the set of all finite values of*

$$\limsup_{k\to+\infty} \frac{1}{q_k\,|q_k\alpha - p_k|},$$

for α irrational, the $\frac{p_k}{q_k}$ being the convergents of α for the Euclid algorithm.

Let us just recall that the Lagrange spectrum is a closed set, its lowest elements are $\sqrt{5}$, then $2\sqrt{2}$, and discrete values (which, including these first two, are called the *Lagrange numbers*) up to a first accumulation point at 3; above 3, its structure is more complicated and not yet fully known, but it contains every real number above a value (which is known to be optimal) near $4,52...$

Note that to get the *Markov spectrum*, we replace the upper limit by a supremum in the above definition; the Markov spectrum will not be used in the present paper.

1.4. **Boshernitzan's ne_n.** In [8] M. Boshernitzan introduced the following quantity:

Definition 1.6. *Let (X_L, S) be a minimal symbolic system. If μ is an S-invariant probability measure, for each natural integer n, we denote by $e_n(\mu)$ the smallest positive frequency of the words of length n of L. If μ is the only invariant probability measure, $e_n(\mu)$ is simply denoted by e_n.*

After partial results in [8] and [26], it was proved in [9] (see also the survey [18]) that whenever, for some invariant probability measure μ, $ne_n(\mu)$ does not tend to 0 when n tends to $+\infty$, then the system (X_L, S) is uniquely ergodic. This sufficient condition for unique ergodicity has been known since [26] as *Boshernitzan's citerion.*

In the present paper, all systems considered are uniquely ergodic, and we consider the quantity ne_n for its own sake. Thus we define

Definition 1.7.

$$\mathcal{B} = \limsup_{n\to+\infty} \frac{1}{ne_n}, \quad \mathcal{B}' = \liminf_{n\to+\infty} \frac{1}{ne_n}.$$

Proposition 1.1. *$\mathcal{B}$ and $\mathcal{B}'$ are invariants of topological conjugacy among uniquely ergodic symbolic dynamical systems.*

Proof
For such a conjugacy ϕ between two symbolic systems, the i-th coordinate of $\phi((x_n)_{n\in\mathbb{N}})$ depends only on $x_i, ..., x_{i+r}$ for a fixed integer r, see for example Lemma 5.1.14 of [24], and the image of the unique invariant probability measure on the first system is the one on the second system. □

A first crude estimate can be given using the complexity function,

Lemma 1.2. $\mathcal{B} \geq \limsup_{n\to+\infty} \frac{p_L(n)}{n}$, $\mathcal{B}' \geq \liminf_{n\to+\infty} \frac{p_L(n)}{n}$.

Proof
As the total measure of the space is 1, we have $e_n \leq \frac{1}{p_L(n)}$. □

This is enough to show that Boshernitzan's criterion is not a necessary condition: there are uniquely ergodic symbolic systems of exponential complexity [22], and thus with $ne_n \to 0$, see also the discussion at the end of Section 4. But of course the above lemma implies that the study of these invariants is interesting only for systems of sub-linear complexity, for which the question of necessity can be asked again.

In view of Theorem 2.5 below, we are led to define the following sets:

Definition 1.8. *For a family of uniquely ergodic symbolic dynamical systems (X_a, S), $a \in \mathcal{F}$, the* upper BL spectrum *is the set of all values of $\mathcal{B}$ taken by the systems in this family, and the* lower BL spectrum *is the set of all values of $\mathcal{B}'$ taken by the systems in this family.*

2. Rotations and the dynamical definition of the Lagrange spectrum

Surely there is nothing new to find about irrational rotations? The computation of $\mathcal{B}$ in Thorem 2.4 below, and the subsequent Theorem 2.5, which was the main motivation for the present paper, were known to P. Hubert and T. Monteil (private communications), but never written to our knowledge. The quantity $\frac{1}{\mathcal{B}'}$ was indeed computed in [12] (see also [5]) as, for irrational rotations, it is equal to another invariant of topological conjugacy, the *covering number by intervals* [12], which involves covering the space by Rokhlin towers; the spectrum of its possible values is the object of a question in [12] and in [10], to which Theorem 2.6 below gives a first (to our knowledge), though belated and partial, answer.

Let $\alpha < \frac{1}{2}$ be an irrational number; the rotations with $\alpha > \frac{1}{2}$ are treated in a similar way and all the results in this section from Theorem 2.4 onwards remain valid; the rotation of angle α, is also the two-interval exchange defined by

$$Tx = \begin{cases} x+\alpha & \text{if } x \in X_1 = [0, 1-\alpha[\\ x-1+\alpha & \text{if } x \in X_2 = [1-\alpha, 1[. \end{cases}$$

With this definition, a rotation admits a natural coding, by the partition of $X = [0,1[$ into X_1 and X_2. Then $L(T)$ has complexity $n+1$ and the trajectories are called *Sturmian sequences.* Irrational rotations are minimal and uniquely ergodic.

To get Theorem 2.4 below, we rely on a computation of both frequencies and lengths of factors of Sturmian sequences, which was done in [4], but which we provide again here by using a different version of the classic Euclid algorithm, making the computations quicker and ready to be generalized. This algorithm is the self-dual induction of [20] in the particular case of two intervals; all what we need to know is contained in the following proposition, which can also be proved directly without difficulty.

Proposition 2.1. *If we build inductively real numbers l_n and r_n and words w_n, M_n, P_n in the following way: $l_1 = \alpha$, $r_1 = 1-2\alpha$, $w_1 = 1$, $M_1 = 1$, $P_1 = 21$. Then*

- *whenever $l_n > r_n$, $l_{n+1} = l_n - r_n$, $r_{n+1} = r_n$, $w_{n+1} = w_nP_n$, $P_{n+1} = P_n$, $M_{n+1} = M_nP_n$;*
- *whenever $r_n > l_n$, $l_{n+1} = l_n$, $r_{n+1} = r_n - l_n$, $w_{n+1} = w_nM_n$, $P_{n+1} = P_nM_n$, $M_{n+1} = M_n$.*

Then the w_n are all the nonempty bispecial words of $L(T)$, w_{n+1} being the shortest bispecial word beginning with w_n. The cylinder $[w_n]$ is the interval $[\alpha - l_n, \alpha + r_n[$; its left and right subintervals separated by α are respectively $T[2w_n]$ and $T[1w_n]$, thus l_n, resp. r_n, is the frequency of the word $2w_n$, resp. $1w_n$. Also, $[w_n1] = [w_nM_n]$ and $[w_n2] = [w_nP_n]$.

The parameters l_n and r_n govern the process; the irrationality of α ensures that always $l_n \neq r_n$. Two fundamental relations come as direct consequences of the above formulas, as we check easily the considered quantities are not changed from n to $n+1$: for all n,

$$l_n|P_n| + r_n|M_n| = 1, \tag{1}$$

$$|M_n| + |P_n| - |w_n| = 2. \tag{2}$$

We define now the *multiplicative* form of this algorithm:

Corollary 2.2. *We define $l_0 = \alpha$, $r_0 = 1-\alpha$, $M_0 = 1$, $P_0 = 2$, w_0 being the empty word. Then we define a_1, a_2, ... such that $r_n > l_n$ for $0 \leq n \leq a_1 - 1$, $r_n < l_n$ for $a_1 \leq n \leq a_1 + a_2 - 1$, and so on. Let $o_k = a_1 + ...a_k$, $o_0 = 0$; let $\alpha_k = r_{o_k}$ and $q_k = |M_{o_k}|$ if k is even, $\alpha_k = l_{o_k}$ and $q_k = |P_{o_k}|$ if k is odd.*

Then for all $k \geq 0$, if $q_{-1} = q_0 = 1$,

$$\alpha_k = a_{k+1}\alpha_{k+1} + \alpha_{k+2}, \quad q_{k+1} = a_{k+1}q_k + q_{k-1}.$$

If k is even $\alpha_{k+1} = l_{o_k}$ and $q_{k-1} = |P_{o_k}|$, while if k is odd $\alpha_{k+1} = r_{o_k}$ and $q_{k-1} = |M_{o_k}|$.

The Euclid continued fraction expansion of α is $[0, a_1+1, a_2, ...]$, and the q_k, $k \geq 0$, are the denominators of the convergents of α.

Proof

The written formulas are straightforward consequences of Proposition 2.1. They imply that $\frac{\alpha_{k+1}}{\alpha_k} = \frac{1}{a_{k+1}+\frac{\alpha_{k+2}}{\alpha_{k+1}}}$ and thus $\frac{\alpha_{k+1}}{\alpha_k}$ has the continued fraction expansion $[0, a_{k+1}, a_{k+2}, ...]$. Going to $k = 0$, $\frac{\alpha_1}{\alpha_0} = \frac{\alpha}{1-\alpha}$ has the continued fraction expansion $[0, a_1, a_2, ...]$, which implies the last assertions. □

The relations (1) and (2) imply $|w_{o_k}| = q_k + q_{k-1} - 2$ and

$$q_k\alpha_k + q_{k-1}\alpha_{k+1} = 1.$$

We can now compute the frequencies as in [4]:

Lemma 2.3. *For $|w_n|+2 \leq s \leq |w_{n+1}|$, the words of length s have three possible frequencies which are l_{n+1}, r_{n+1}, and $l_{n+1}+r_{n+1}$; for $s = |w_n|+1$, the words of length s have two possible frequencies which are l_n and r_n.*

Proof

As in [4], we build the Rauzy graphs. If w is not right special, w' is not left special, and $w \to w'$, then w and w' have the same frequency.

In the case of Sturmian sequences, for each s, there is one right special word of length s and there is one left special word G_s of length s; there are at most three frequencies, which are for example those of G_s, $1G_s$ and $2G_s$, and these are exactly $l_m + r_m$, l_m and r_m for the smallest m such that $|w_m| \geq s$.

But we need to check whether each of these frequencies is indeed the frequency of some word. We take an n such that $r_n > l_n$, and let p be the length of w_n. The possible frequencies of words of length p are l_n, r_n, $l_n + r_n$, each one is indeed the frequency of a word of length p, as we take respectively $2w_n$ deprived of its last letter, $1w_n$ deprived of its last letter, and w_n. Then the words of length $p+1$ have possible frequencies l_{n+1}, r_{n+1}, $l_{n+1} + r_{n+1}$, which are respectively l_n, $r_n - l_n$, r_n; each of the two old frequencies l_n, r_n is the frequency of a word of length $p+1$, as we take respectively $2w_n$ and $1w_n$, but the new one $r_n - l_n$ is not the frequency of any word of length $p+1$. The words of length $p+2$ have possible frequencies l_{n+1}, r_{n+1}, $l_{n+1} + r_{n+1}$, and each one is the frequency of a word of length $p+2$, as we take respectively $2w_n1$, $1w_n1$, and the prefix of length $p+2$ of w_nM_n; and this remains true (by extending the considered words to the right following w_nM_n) for $p+3$,... until we reach the length of w_{n+1}. And the reasoning is similar if $l_n > r_n$. □

Theorem 2.4. *For a rotation of irrational angle* $\alpha = [0, b_1, .., .b_n, ...]$*, if we define* $v_k = [0, b_k, b_{k-1}, ...b_1]$ *and* $t_k = [0, b_{k+1}, b_{k+2}, ...]$ *then*

$$\mathcal{B} = \limsup_{k\to+\infty}\left(\frac{1}{v_k} + t_k\right) = \limsup_{k\to+\infty}\left(b_k + v_{k-1} + t_k\right), \quad \mathcal{B}' = \liminf_{k\to+\infty}\left(1 + t_k v_k\right).$$

Proof

Suppose k is even. Then for $o_k \le m \le o_{k+1} - 1$, $l_m = l_{o_k}$ and $l_m < r_m$; thus the smallest frequency of a word of length n is l_{o_k} for $|w_{o_k-1}| + 2 \le n \le |w_{o_{k+1}-1}| + 1$.

We know that $l_{o_k} = \alpha_{k+1}$ and $|w_{o_k}| = q_k + q_{k-1} - 2$, and we check $|w_{o_k-1}| = q_k - 2$. Thus the minimal value of ne_n for $|w_{o_k-1}| + 2 \le n \le |w_{o_{k+1}-1}| + 1$ is $q_k\alpha_{k+1}$; the maximal value of ne_n for $|w_{o_k-1}| + 2 \le n \le |w_{o_{k+1}-1}| + 1$ is $q_{k+1}\alpha_{k+1}$; all this is still true if k is odd.

In Corollary 2.2, we identify b_1 with $a_1 + 1$ and b_i with a_i for $i \ge 2$. Thus we get $\frac{\alpha_{k+1}}{\alpha_k} = t_k$, while by construction of the q_k we get $\frac{q_{k-1}}{q_k} = v_k$.

Then $\limsup_{n\to\infty} \frac{1}{ne_n} = \limsup_{k\to\infty} \frac{1}{q_k\alpha_{k+1}} = \limsup_{k\to\infty} \frac{q_{k+1}\alpha_{k+1} + q_k\alpha_{k+2}}{q_k\alpha_{k+1}} = \limsup_{k\to\infty}(\frac{q_{k+1}}{q_k} + \frac{\alpha_{k+2}}{\alpha_{k+1}})$, which is the first formula at stage $k+1$.

Similarly, $\liminf_{n\to\infty} \frac{1}{ne_n} = \liminf_{k\to\infty} \frac{1}{q_{k+1}\alpha_{k+1}} = \liminf_{k\to\infty} \frac{q_{k+1}\alpha_{k+1} + q_k\alpha_{k+2}}{q_{k+1}\alpha_{k+1}} = \liminf_{k\to\infty}(1 + \frac{\alpha_{k+2}}{\alpha_{k+1}}\frac{q_k}{q_{k+1}})$, which is the second formula at stage $k+1$. □

Theorem 2.5. *The upper BL spectrum of the family of rotations is the union of the Lagrange spectrum and* $+\infty$.

Proof

We check that $p_k\alpha_k + p_{k-1}\alpha_{k+1} = \alpha$ and $|p_kq_{k-1} - p_{k-1}q_k| = 1$, thus $\liminf_{k\to\infty} q_k|q_k\alpha - p_k|$ is equal to $\liminf_{k\to\infty} q_k\alpha_{k+1}$. □

As for the lower LB spectrum, it seems to have never been studied to our knowledge, and its study looks to be of the same level of difficulty as for the Lagrange spectrum. We give now some of the first results about it. Note that $\mathcal{B}'$ is not the lower limit of $\frac{1}{q_k|q_k\alpha - p_k|}$ and thus is not directly linked to the quality of the approximation of α by rationals.

Theorem 2.6. *The lower BL spectrum of the family of rotations has* 1 *as its smallest element, with* $\mathcal{B}' = 1$ *if and only if the angle has unbounded partial quotients. It is a closed set.*
Its two largest elements are $\frac{5-\sqrt{5}}{2} = 1,38196...$ *and* $3 - \sqrt{3} = 1,26794...$*, and there is no other element above* $\frac{5}{4}$.
It contains an accumulation point equal to $\sqrt{5} - 1 = 1,2360...$
It contains the interval $[1, 1 + \frac{4}{83+18\sqrt{2}} = 1,03688...]$.

Proof

The first two assertions are straightforward consequences of the formula giving $\mathcal{B}'$. The spectrum is closed by the standard reasoning of [14], Chapter 1, Corollary to Lemma 6, as every value greater than one is reached by an angle with bounded partial quotients, and for any sequence of angles such that the corresponding $\mathcal{B}'$ converge to a number different from one, we can take the partial quotients uniformly bounded.

Let $\alpha = [0, b_1, .., .b_n, ...]$ be the angle of a rotation.

Suppose that there are infinitely many $b_k \geq 4$; then for these k, $t_k v_k \leq \frac{1}{4}$.
If there are only 1, 2 or 3 with infinitely many 3, we have $t_k \leq [0,(13)^\omega]$. If $b_k = 3$, we have $v_k \leq [0, 3(31)^\omega]$ and we get $t_k v_k \leq \frac{\sqrt{21}-3}{\sqrt{21}+15} = 0,2424...$
If there are only 1 and 2 with infinitely many 22, if $b_k = b_{k+1} = 2$ we get $t_k \leq [0, 2(21)^\omega]$ and $v_k \leq [0, 2(21)^\omega]$, thus $t_k v_k \leq \frac{12-6\sqrt{3}}{9} = 0,1786...$
If there are only 1 and 2 with no 22, infinitely many 2 and infinitely many 11, if $b_k = 2$, $b_{k+1} = b_{k+2} = 1$ we get $v_k \leq [0, 21(12)^\omega]$, $t_k \leq [0, 11(12)^\omega]$ and $t_k v_k \leq \frac{6-\sqrt{3}}{11}\frac{3-\sqrt{3}}{2} = 0,2459...$

Thus the two highest values are reached for the angles $[0,(21)^\omega]$ and $[0,(1)^\omega]$, and the above estimates give a bound on the gap below them.

If we take $\alpha = [0,(1^j 2)^\omega]$, for a $j \geq 3$, the smallest value of $t_n v_n$ is reached when $b_k = 2$ or $b_{k+1} = 2$, and when j tends to infinity, this value tends to $[0,(1)^\omega][0,2(1)^\omega] = \sqrt{5}-2$. Thus we get an accumulation point, approached from above if j is odd.

To prove the last assertion, we use Theorem 3.2 of [23]: any real number $s > 1$ can be written $s = (r + [0, a_1, a_2, ...])(r' + [0, a'_1, a'_2, ...])$ with all the a_i and a'_i taking values 1, 2, 3 or 4. Moreover, from the proof of this theorem we get ([23] p. 974) that if s is at least $n^2 + (\sqrt{2}-1)n + \frac{3-2\sqrt{2}}{4}$ we can take both r and r' greater or equal to n. Thus if s is at least $\frac{83+18\sqrt{2}}{4}$, we write s in that way, with r and r' at least 5. We choose now a sequence $k_n \to +\infty$, and take the rotation of angle $[0, a_{k_1}, a_{k_1-1}, .., a_1, r, r', a'_1, ...a'_{k_1}, a_{k_2}, a_{k_2-1}, .., a_1, r, r', a'_1, ...a'_{k_2}, ...]$; then the minimal values of $t_k v_k$ are taken when $a_k = r$, and their lower limit is exactly $\frac{1}{s}$. □

The third highest number in this spectrum is $\frac{16-4\sqrt{6}}{5} = 1,2404...$, as can be seen with longer computations; the point $\sqrt{5}-1$ is the highest accumulation point, but to prove it requires a machinery similar to the one used to prove Theorem 5 in Chapter 1 of [14].

3. Three-interval exchanges

3.1. **The transformations.**

Definition 3.1. *Given two numbers* $0 < \alpha$, $0 < \beta$ *with* $\alpha + \beta < 1$, *we define a* three-interval exchange *on* $X = [0, 1[$ *by*

$$Tx = \begin{cases} x + 1 - \alpha & \text{if } x \in X_1 = [0, \alpha[\\ x + 1 - 2\alpha - \beta & \text{if } x \in X_2 = [\alpha, \alpha+\beta[\\ x - \alpha - \beta & \text{if } x \in X_3 = [\alpha+\beta, 1[. \end{cases}$$

Throughout this section, we ask that α and β satisfy the *i.d.o.c condition* of Keane, which means in that case that they do not satisfy any rational relation of the forms $p\alpha + q\beta = p - q$, $p\alpha + q\beta = p - q + 1$, or $p\alpha + q\beta = p - q - 1$, for p and q integers.

The points α and $\alpha + \beta$ are the discontinuities of T, while $\beta_1 = 1 - \alpha - \beta$ and $\beta_2 = 1 - \alpha$ are the discontinuities of T^{-1}. The i.d.o.c. condition ensures that the negative orbits of the discontinuities of T are infinite and have an empty intersection (it is its original definition; see [16] for the equivalence with the one stated here).

A three-interval exchange admits a natural coding, by the partition of X into X_1, X_2, X_3. Under the i.d.o.c. condition, (X, T) is minimal and uniquely ergodic and $L(T)$ has complexity $2n + 1$.

We recall that the *induced*, or first return, map S of a transformation T on a set E is defined on E by $Sx = T^{g(x)}x$, where $g(x)$ is the smallest integer $r > 0$ such that $T^r x$ is in E (a finite $g(x)$ does indeed exist in all cases occurring in the present paper).

A three-interval exchange defined as above is always the induced map of the (irrational) rotation of angle $\frac{1-\alpha}{1+\beta}$ on the interval $[0, \frac{1}{1+\beta}[$.

Throughout this section, we add the conditions $0 < \alpha < \frac{1}{2}$, and $2\alpha + \beta > 1$; they ensure that the induction process described below does not have an irregular behaviour in the early stages: as is shown in [20], their absence modifies only a finite number of stages, and all the results in this section from Theorem 3.10 onwards remain valid without these extra conditions.

3.2. **The self-dual induction for 3 intervals.** We state now some results from [20].

Definition 3.2. *We define an operation called* self-dual induction*, which builds numbers $l_{i,n}$ and $r_{i,n}$, $i = 1, 2$, through an infinite sequences of* states *in the following way:*

State I *is defined by the relation $r_{1,n} = r_{2,n}$.*

In Substate Ia $l_{1,n} > r_{1,n}$, $l_{2,n} > r_{2,n}$. *Then for $i = 1, 2$ we put $l_{i,n+1} = l_{i,n} - r_{i,n}$, $r_{i,n+1} = r_{i,n}$. For $n + 1$ we are again in state I.*

In Substate Ib $l_{1,n} < r_{1,n}$, $l_{2,n} > r_{2,n}$. *Then $l_{1.n+1} = l_{1,n}$, $r_{1,n+1} = r_{1,n}$, $l_{2,n+1} = l_{2,n} - r_{2,n}$, $r_{2,n+1} = r_{2,n}$. For $n + 1$ we are again in state I.*

In Substate Ic $l_{1,n} > r_{1,n}$, $l_{2,n} < r_{2,n}$. *This is deduced from Ib by exchanging* 1 *and* 2*, and for $n + 1$ we are again in state I.*

In Substate Id $l_{1,n} < r_{1,n}$, $l_{2,n} < r_{2,n}$. *Then for $i = 1, 2$, $l_{i,n+1} = l_{i,n}$, $r_{i,n+1} = r_{i,n} - l_{i,n}$. For $n + 1$ we are in state II described just below.*

State II *is defined by the relation $l_{1,n} + r_{1,n} = l_{2,n} + r_{2,n}$. Note that in this state $l_{1,n} > r_{2,n}$ if and only if $l_{2,n} > r_{1,n}$.*

In Substate IIa $l_{1,n} > r_{2,n}$, $l_{2,n} > r_{1,n}$. *Then for $i = 1, 2$, $l_{i.n+1} = l_{i,n} - r_{3-i,n}$, $r_{i,n+1} = r_{i,n}$. For $n + 1$ we are in state III described below.*

In Substate IIb $l_{1,n} < r_{2,n}$, $l_{2,n} < r_{1,n}$. *Then for $i = 1, 2$, $r_{i.n+1} = r_{i,n} - l_{3-i,n}$, $l_{i,n+1} = l_{i,n}$. For $n + 1$ we are in state I.*

State III *is symmetrical to state I, with left and right exchanged, and the relation $l_{1,n} = l_{2,n}$; there are four substates, $IIIa$ to $IIId$, and the induction goes either to state III or to state II.*

Proposition 3.1. *For a given three-interval exchange, we build inductively real numbers $l_{i,n}$ and $r_{i,n}$ and words $w_{i,n}$, $M_{i,n}$, $P_{i,n}$, $i = 1, 2$, in the following way: $l_{1,1} = 1 - \alpha - \beta$, $r_{1,1} = 2\alpha + \beta - 1 = r_{2,1}$, $l_{2,1} = 1 - 2\alpha$, $w_{i,1} = i$ and $M_{i,1} = i$ for $i = 1, 2$, $P_{1,1} = 31$, $P_{2,1} = 2$.*

$l_{i,n}$ and $r_{i,n}$ are built by Definition 3.2, starting from $n = 1$, for which we are in State I.

For each $n > 0$; let $s_n(1) = 1$, $s_n(2) = 2$ if for n we are in I or III, $s_n(1) = 2$, $s_n(2) = 1$ if for n we are in II; $p_n(1) = 1$, $p_n(2) = 2$ if for n we are in I, $p_n(1) = 2$, $p_n(2) = 1$ if for n we are in II or III; $m_n(1) = 1$, $m_n(2) = 2$ if for n we are in III, $m_n(1) = 2$, $m_n(2) = 1$ if for n we are in I or II. Then for $1 \le i \le 2$

- *whenever* $l_{i,n+1} < l_{i,n}$, *then* $w_{i,n+1} = w_{i,n} P_{s_n p_n(i),n}$, $P_{i,n+1} = P_{i,n}$, $M_{i,n+1} = M_{p_n(i),n} P_{i,n}$;
- *whenever* $r_{i,n+1} < r_{i,n}$, *then* $w_{i,n+1} = w_{i,n} M_{s_n m_n(i),n}$, $P_{i,n+1} = P_{m_n(i),n} M_{i,n}$, $M_{i,n+1} = M_{i,n}$;
- *otherwise* $w_{i,n+1} = w_{i,n}$, $P_{i,n+1} = P_{i,n},$, $M_{i,n+1} = M_{i,n}$.

Then the $w_{i,n}$, $n = 1, 2$, *are all the nonempty bispecial words of* $L(T)$; *either* $w_{i,n+1} = w_{i,n}$ *or* $w_{i,n+1}$ *is the shortest bispecial word beginning with* $w_{i,n}$. *The cylinder* $[w_{i,n}]$ *is the interval* $[\beta_i - l_{i,n}, \beta_i + r_{i,n}[$; *its left and right subintervals separated by* β_i *are respectively* $T[bw_{i,n}]$ *and* $T[aw_{i,n}]$ *if* $bw_{i,n}$ *and* $aw_{i,n}$, $a < b$, *are the two left extensions of* $w_{i,n}$; *thus indeed* $l_{1,n}$, $r_{1,n}$, $l_{2,n}$, $r_{2,n}$ *are respectively the frequencies of the words* $3w_{1,n}$, $2w_{1,n}$, $2w_{2,n}$, $1w_{2,n}$. *Also,* $[w_{i,n} M_{s_n m_n(i),n}] = [w_{i,n} c]$ *and* $[w_{i,n} P_{s_n p_n(i),n}] = [w_{i,n} d]$, *if* $w_{i,n} c$ *and* $w_{i,n} d$, $d > c$, *are the two right extensions of* $w_{i,n}$.

A three-interval exchange defines an infinite sequence of states labelled Ia *to* Id, IIa, IIb, $IIIa$ *to* $IIId$, *following the rules of Definition 3.2, and such that each one of the four parameters* $l_{1,n}$, $r_{1,n}$, $l_{2,n}$, $r_{2,n}$ *is modified infinitely often. Conversely, every such sequence of states defines a three-interval exchange which generates it as described above.*

Note that the i.d.o.c. condition ensures that we dont have $l_{1,n} = r_{1,n}$ or similar equalities. We check from the formulas that the following bilinear form is invariant by the induction, and thus for all n

$$l_{1,n}|P_{1,n}| + r_{1,n}|M_{1,n}| + l_{2,n}|P_{2,n}| + r_{2,n}|M_{2,n}| = 1. \tag{3}$$

The other relations are more complicated than in the case of rotations, we state them now but they are a straightforward consequence of Corollary 3.2 below:

- whenever for n we are in II,

$$|P_{2,n}| + |M_{1,n}| = |M_{2,n}| + |P_{1,n}| + 1 = |w_{2,n}| + 2 = |w_{1,n}| + 2; \tag{4}$$

- whenever for n we are in I,

$$|P_{1,n}| + |M_{1,n}| = |w_{1,n}| + 2, \quad |P_{2,n}| + |M_{2,n}| = |w_{2,n}| + 1, \quad |P_{1,n}| = |P_{2,n}| + 1;$$

- whenever for n we are in III,

$$|P_{1,n}| + |M_{1,n}| = |w_{1,n}| + 1, \quad |P_{2,n}| + |M_{2,n}| = |w_{2,n}| + 2, \quad |M_{2,n}| = |M_{1,n}| + 1.$$

We use now the self-dual induction to retrieve the frequencies and the lengths of the bispecial words, and this involves a multiplicative form of the algorithm; this has been done in [15] [16] [17] but in a form which is more complicated and less explicit (the frequencies of words are somewhat hidden) than the one which we deduce now from [20]; the results in this section are new, but could be deduced from [15] and [16], see [19] for the equivalence between the two forms of the algorithm. Thus the present paper is independent of all the previous ones except [20], and all the information we need from the latter are in Definition 3.2 and Proposition 3.1 above.

Corollary 3.2. *For a given 3-iet, we define* o_k, $k = 1, 2...$ *to be the sequence of* $n \geq 1$ *such that for* n *we are in state* II; *we define also* $o_0 = 0$. *Then for* $k \geq 1$, *we define two positive integers,* n_k, *resp.* m_k, *as the number of* $o_{k-1} < n < o_k$ *such that* $w_{1,n+1} \neq w_{1,n}$, *resp.* $w_{2,n+1} \neq w_{2,n}$. *Finally we define a sequence* η_k, *which is* $-$, *resp.* $+$, *whenever for* $n = o_k - 1$ *we are in state* I, *resp.* III.

We define also parameters for $n = 0$, *namely, we are in state* II, $l_{1,0} = 1-\alpha-\beta$, $r_{1,0} = \beta$, $l_{2,0} = 1-2\alpha$, $r_{2,0} = \alpha$, $w_{1,0}$, $w_{2,0}$ *and* $P_{1,0}$ *are the empty word,* $P_{2,0} = 3$, $M_{1,0} = 1$, $M_{2,0} = 2$, $\eta_0 = +$.

Then $o_0 = 0$, $\eta_1 = -$. *In the case* $\eta_{k+1} = -$, *for* $0 \le a \le n_{k+1}-1$, $0 \le b \le m_{k+1}-1$,

$$r_{1,o_k+1+a} = r_{1,o_k} - l_{2,o_k} = r_{2,o_k} - l_{1,o_k}, \quad l_{1,o_k+1+a} = l_{1,o_k} - a((r_{1,o_k} - l_{2,o_k}),$$

$$r_{2,o_k+1+b} = r_{1,o_k} - l_{2,o_k} = r_{2,o_k} - l_{1,o_k}, \quad l_{2,o_k+1+b} = l_{2,o_k} - b(r_{1,o_k} - l_{2,o_k}),$$

$$w_{1,o_k+1+a} = w_{1,o_k} M_{1,o_k} (P_{2,o_k} M_{1,o_k})^a,$$

$$w_{2,o_k+1+b} = w_{2,o_k} M_{2,o_k} (P_{1,o_k} M_{2,o_k})^b,$$

$$M_{1,o_k+1+a} = M_{1,o_k} (P_{2,o_k} M_{1,o_k})^a, \quad P_{1,o_k+1+a} = P_{2,o_k} M_{1,o_k};$$

$$M_{2,o_k+1+b} = M_{2,o_k} (P_{1,o_k} M_{2,o_k})^b, \quad P_{2,o_k+1+b} = P_{1,o_k} M_{2,o_k}.$$

Then $o_{k+1} = \max\{o_k + 1 + n_{k+1}, o_k + 1 + m_{k+1}\}$; *between* $\min\{o_k + n_{k+1}, o_k + m_{k+1}\}$ *and* $\max\{o_k + n_{k+1}, o_k + m_{k+1}\}$ *one of the intervals is not modified, and*

$$r_{1,o_{k+1}} = n_{k+1}(r_{1,o_k} - l_{2,o_k}) - l_{1,o_k}, \quad l_{1,o_{k+1}} = l_{1,o_k} - (n_{k+1}-1)((r_{1,o_k} - l_{2,o_k}),$$

$$r_{2,o_{k+1}} = m_{k+1}(r_{1,o_k} - l_{2,o_k}) - l_{2,o_k}, \quad l_{2,o_{k+1}} = l_{2,o_k} - (m_{k+1}-1)((r_{1,o_k} - l_{2,o_k}),$$

$$w_{1,o_{k+1}} = w_{1,o_k} M_{1,o_k} (P_{2,o_k} M_{1,o_k})^{n_{k+1}-1} M_{2,o_k} (P_{1,o_k} M_{2,o_k})^{m_{k+1}-1},$$

$$w_{2,o_{k+1}} = w_{2,o_k} M_{2,o_k} (P_{1,o_k} M_{2,o_k})^{m_{k+1}-1} M_{1,o_k} (P_{2,o_k} M_{1,o_k})^{n_{k+1}-1},$$

$$P_{1,o_{k+1}} = P_{1,o_k} M_{2,o_k} M_{1,o_k} (P_{2,o_k} M_{1,o_k})^{n_{k+1}-1}, \quad M_{1,o_{k+1}} = M_{1,o_k} (P_{2,o_k,1} M_{1,o_k})^{n_{k+1}-1},$$

$$P_{2,o_{k+1}} = P_{2,o_k} M_{1,o_k} M_{2,o_k} (P_{1,o_k} M_{2,o_k})^{m_{k+1}-1}, \quad M_{2,o_{k+1}} = M_{2,o_k} (P_{1,o_k} M_{2,o_k})^{m_{k+1}-1}.$$

We get the case $\eta_{k+1} = +$ *by exchanging the* l *and* r, *or the* M *and* P, *in both sides of all the above equalities.*

A three-interval exchange defines an infinite sequence (n_k, m_k, η_k), $k \ge 1$, $n_k \ge 1$, $m_k \ge 1$, $\eta_k = \pm$, *such that we do not have ultimately* $n_k = 1$ *and* η_k *constant, or* $m_k = 1$ *and* η_k *constant. Conversely, every such sequence defines a three-interval exchange which generates it as described above.*

Proof

Through all the substates of state I we have $r_{i,n+1} = r_{i,n}$, and the $l_{i,n}$ are decreased by a fixed quantity, so after a finite number of steps we are in Id and proceed to II, and similarly throughout stage III. Thus we are infinitely many times in state II and we can define o_k; n_k and m_k are at least one as for $n = o_k + 1$ we are in state I or III; we check that our parameters for $n = 0$ are compatible with the ones already defined for $n = 1$. Then we apply recursively Definition 3.2 and Proposition 3.1. □

In the sequel, for a given three-interval exchange T, we shall use either the sequence (n_k, m_k, η_k) or $(n_k, m_k, -\eta_k\eta_{k+1})$, which we both call the *expansion* of T; an expansion is *admissible* if we do not have ultimately $n_k = 1$ and η_k constant, or $m_k = 1$ and η_k constant. Note that [15] [16] [17] [19] use the expansion $(n_k, m_k, \epsilon_{k+1})$ where $\epsilon_{k+1} = \eta_k\eta_{k+1}$.

We make now some further computations on the formulas above.

Lemma 3.3. *For all $k \geq 1$,*

$$\begin{pmatrix} r_{1,o_{k-1}} + r_{2,o_{k-1}} \\ r_{1,o_{k-1}} - r_{2,o_{k-1}} \\ l_{1,o_{k-1}} + r_{1,o_{k-1}} \end{pmatrix} = \begin{pmatrix} \eta_k & 0 & m_k + n_k - 2\eta_k \\ 0 & 1 & (n_k - m_k)\eta_k \\ \eta_k & 0 & m_k + n_k - \eta_k \end{pmatrix} \begin{pmatrix} r_{1,o_k} + r_{2,o_k} \\ r_{1,o_k} - r_{2,o_k} \\ l_{1,o_k} + r_{1,o_k} \end{pmatrix}. \tag{5}$$

Let $\Delta_k = l_{1,o_k} + r_{1,o_k} = l_{2,o_k} + r_{2,o_k}$, $t_k = \frac{\Delta_{k+1}}{\Delta_k}$. Then

$$r_{1,o_k} + r_{2,o_k} - \Delta_k = -\eta_{k+1}\Delta_{k+1}, \tag{6}$$

$$t_k = [0, (m_{k+1} + n_{k+1}) * -\eta_{k+1}\eta_{k+2}, (m_{k+2} + n_{k+2}) * -\eta_{k+2}\eta_{k+3}, ...] \tag{7}$$

$$\frac{1-\alpha}{1+\beta} = [0, 2*+, (m_1 + n_1) * -\eta_1\eta_2, (m_2 + n_2) * -\eta_2\eta_3, ...] \tag{8}$$

Proof
(5) comes directly from the formulas in Corollary 3.2. These imply also (6), and $t_k = -\eta_{k+1}\frac{r_{1,o_k}+r_{2,o_k}-\Delta_k}{\Delta_k}$, thus we get $t_k = \frac{1}{m_{k+1}+n_{k+1}-\frac{\eta_{k+1}\eta_{k+2}}{t_{k+1}}}$, and t_k is given by the semi-regular continued fraction expansion (7). Going to $k = 0$ we get a semi-regular continued fraction expansion of $\frac{\alpha+\beta}{1-\alpha}$, which can be put in the form (8). □

Lemma 3.4. *For all $k \geq 1$,*

$$\begin{pmatrix} |P_{2,o_k}| + |P_{1,o_k}| \\ |P_{1,o_k}| - |P_{2,o_k}| \\ |M_{2,o_k}| + |P_{1,o_k}| \end{pmatrix} = V_k + \begin{pmatrix} \eta_k & 0 & m_k + n_k - 2\eta_k \\ 0 & 1 & n_k - m_k \\ \eta_k & 0 & m_k + n_k - \eta_k \end{pmatrix} \begin{pmatrix} |P_{2,o_{k-1}}| + |P_{1,o_{k-1}}| \\ |P_{1,o_{k-1}}| - |P_{2,o_{k-1}}| \\ |M_{2,o_{k-1}}| + |P_{1,o_{k-1}}| \end{pmatrix}. \tag{9}$$

where

$$V_k = \begin{pmatrix} n_k + 1 \\ n_k - 1 \\ n_k \end{pmatrix} \quad \text{if } \eta_k = -, \quad V_k = \begin{pmatrix} m_k - 1 \\ 1 - m_k \\ m_k \end{pmatrix} \quad \text{if } \eta_k = +.$$

We define $z_k = |M_{2,o_k}| + |P_{1,o_k}|$, $v_k = \frac{q_{k-1}}{q_k}$, where

$$q_{k+1} = (m_{k+1} + n_{k+1})q_k - \eta_k\eta_{k+1}q_{k-1},$$

$q_{-1} = 0$, $q_0 = 1$. Then

$$|M_{2,o_k}| + |M_{1,o_k}| = z_k + \eta_k z_{k-1} + 1 + \eta_k. \tag{10}$$

$$\lim_{k\to+\infty} (z_k\Delta_k - \eta_k\eta_{k+1}z_{k-1}\Delta_{k+1}) = 1. \tag{11}$$

$$v_k = [0, (m_k + n_k) * -\eta_{k-1}\eta_k, (m_{k-1} + n_{k-1}) * -\eta_{k-2}\eta_{k-1}, ..., (m_2 + n_2) * -\eta_1\eta_2, m_1 + n_1] \tag{12}$$

$$\lim_{k\to+\infty} \left(\frac{z_{k-1}}{z_k} - v_k\right) = 0. \tag{13}$$

Proof
(9) comes directly from the formulas in Corollary 3.2. This implies in turn $|P_{2,o_k}| + |P_{1,o_k}| = z_k - \eta_k z_{k-1} - \eta_k$ and thus (10).

We write now the relation (3) for $n = o_k + 1$. Suppose $\eta_{k+1} = -1$. We notice that $r_{1,o_k+1} = r_{2,o_k+1} = r_{1,o_k} - l_{2,o_k} = r_{2,o_k} - l_{1,o_k}$ is also equal to Δ_{k+1}, and thus $\Delta_{k+1}(|M_{2,o_k}| +$

$|M_{1,o_k}|)+l_{1,o_k}(|P_{2,o_k}|+|M_{1,o_k}|)+l_{2,o_k}(|M_{2,o_k}|+|P_{1,o_k}|)=1$. We recall that $|M_{2,o_k}|+|P_{1,o_k}|=z_k$ and $|P_{2,o_k}|+|M_{1,o_k}|=z_k+1$, while $|M_{2,o_k}|+|M_{1,o_k}|$ is given by (10). Because of (6), we get finally that $z_k\Delta_k+\eta_k z_{k-1}\Delta_{k+1}+(1+\eta_k)\Delta_{k+1}+l_{1,o_k}=1$. Similar computations when $\eta_{k+1}=+1$ give $z_k\Delta_k-\eta_k z_{k-1}\Delta_{k+1}-\eta_k\Delta_{k+1}+r_{2,o_k}=1$. In both cases we get (11).

(12) is a straightforward consequence of the definition of q_k.

We show now that $\frac{z_{k-1}}{z_k}-\frac{q_{k-1}}{q_k}$ tends to 0 when k tends to $+\infty$. Indeed, from the matrix equality with B_k and (10) we get $z_{k+1}=(m_{k+1}+n_{k+1})z_k-\eta_k\eta_{k+1}z_{k-1}-\eta_k\eta_{k+1}+\theta_{k+1}$, where $\theta_k=n_k$ if $\eta_k=-1$, $\theta_k=m_k$ if $\eta_k=+1$. If $\Phi_k=q_kz_{k-1}-q_{k_1}z_k$, then $\Phi_{k+1}=-\eta_k\eta_{k+1}\Phi_k-q_k(\theta_{k+1}+\eta_k\eta_{k+1})$; we see that Φ_k grows more slowly than q_k, and this implies (13) as z_k tends to infinity with k. □

Note that the matrices in Lemmas 3.3 and 3.4 are the same except for one entry, the self-duality of the induction works but not as perfectly as for the rotations.

Though this will not be used in the sequel, the q_k are indeed present in the computation of the lengths of the intervals, as we check that the third line of the matrix $A_1...A_k$ is made with the entries η_kq_{k-1}, 0, $q_k-\eta_kq_{k-1}$, and this implies that $q_k\Delta_k-\eta_k\eta_{k+1}q_{k-1}\Delta_{k+1}=1-\alpha$. Thus from (11), (13), and the last equality, we deduce that $\frac{q_k}{z_k}\to 1-\alpha$ when k tends to infinity.

From now on, when we study quantities involving large k, we shall always discard higher order terms such as $|w_{1,o_k}|-z_k$ or $z_k\Delta_k-z_{k-1}\Delta_{k+1}-1$, and replace $\frac{z_{k-1}}{z_k}$ by v_k.

3.3. **General formulas for $\mathcal{B}$ and $\mathcal{B}'$.**

Lemma 3.5. *The frequencies of words of length n take at most six different values, which are $l_{i,q(i,n)}$, $r_{i,q(i,n)}$, and $l_{i,q(i,n)}+r_{i,q(i,n)}$, $i=1,2$, where $w_{i,q(i,n)}$ is the shortest bispecial word (in the families built by the self-dual induction) of length at least n.*

Proof

We check that if $w=w_1...w_r$ is in $L(T)$, so is $\overline{w}=w_r...w_1$, with the same frequency as w: this is done by induction on the length of w, as the frequencies of the words of length n are deduced from the frequencies of words of length $n-1$ and the expression of the words of length n, see for example [8] or Proposition 4 of [21].

Then we build the Rauzy graphs, as was done in [8] without naming them. Given w, there exists at least one (left or right) special word linked by arrows with w, otherwise there is a periodic orbit and this contradicts the i.d.o.c. condition; thus w has the same frequency as v, where v is a word with n letters, and either v is left special, or $v\to v'$ where v' is left special, or v is right special, or $v'\to v$ where v' is right special; by looking also at $\overline{w}$, we can suppose v is left special or $v\to v'$ where v' is left special.

There are two left special words with n letters, $v_{i,n}$ beginning with i; let $L_{i,n}$ be the frequency of $bv_{i,n}$, $R_{i,n}$ the frequency of $av_{i,n}$, for the two left extensions of $v_{i,n}$, $b>a$. Then the frequency of $v_{i,n}$ is $L_{i,n}+R_{i,n}$, and the frequency of a v such that $v\to v_{i,n}$ is either $L_{i,n}$ or $R_{i,n}$. Moreover, let $W_{i,n}$ be the shortest bispecial word having $v_{i,n}$ as a prefix, which is also the shortest bispecial word beginning with i and with length at least n; then $W_{i,n}$ is $w_{i,q}$ for $q=q(i,n)$, $L_{i,n}=l_{i,q}$ and $R_{i,n}=r_{i,q}$. □

Lemma 3.5 was proved in [20] and can also be deduced from [8], we reprove it for sake of completeness.

Lemma 3.6. *Suppose $r_{i,n+1} < r_{i,n}$; then each of the frequencies $l_{i,n+1}$ and $l_{i,n+1} + r_{i,n+1}$ is the frequency of a word of length s for every $|w_{i,n}| + 1 \leq s \leq |w_{i,n+1}|$; the frequency $r_{i,n+1}$ is the frequency of a word of length s for every $|w_{i,}| + 2 \leq s \leq |w_{i,n+1}|$, but not for $s = |w_{i,n}| + 1$. The same is true with l and r exchanged.*

Proof

This is the same as the last part of the proof of Lemma 2.3, mutatis mutandis: in particular, while the word w_n could be preceded by two letters $2 > 1$, and followed by two letters $1 < 2$, here the word $w_{i,n}$ can be preceded by two letters $b > a$, and followed by two letters $c < d$, and $1w_n2$ is replaced by $aw_{i,n}d$ and so on. □

Proposition 3.7. *Let T be a three-interval exchange with expansion $(n_k, m_k, -\eta_k\eta_{k+1})$; let t_k be as in Lemma 3.3, v_k as in Lemma 3.4; let*

$$t_{k,l} = \frac{l_{2,o_k}}{l_{1,o_k}} \quad t_{k,r} = \frac{r_{2,o_k}}{r_{1,o_k}}, \quad v_{k,M} = \frac{(m_k - 1)z_{k-1} + |M_{2,o_{k-1}}|}{(n_k - 1)z_{k-1} + |M_{1,o_{k-1}}|}, \quad v'_{k,M} \frac{m_k z_{k-1} + |M_{2,o_{k-1}}|}{n_k z_{k-1} + |M_{1,o_{k-1}}|},$$

$$B_k = \frac{1}{1 - \eta_k\eta_{k+1}t_k} + \frac{v_k}{1 - v_k}, \quad B'_k = \frac{1}{1 + \eta_k\eta_{k+1}t_k} - \frac{v_k}{v_k + 1}.$$

Then $\mathcal{B}$ is the largest of the upper limits, when $k \to \infty$, of the following quantities

$$\frac{1}{v_k} - \eta_k\eta_{k+1}t_k, \tag{14}$$

$$(1 + t_{k,l})\,(1 + v_{k,M})\,B_k, \tag{15}$$

$$\left(1 + \frac{1}{t_{k,l}}\right)\left(1 + \frac{1}{v_{k,M}}\right)B_k, \tag{16}$$

$$(1 + t_{k,r})\left(1 + v'_{k,M}\right)B'_k, \tag{17}$$

$$\left(1 + \frac{1}{t_{k,r}}\right)\left(1 + \frac{1}{v'_{k,M}}\right)B'_k, \tag{18}$$

where in formulas (15) to (18), when $\eta_k = +$, $t_{k,l}$ and $t_{k,r}$ have to be exchanged, and $v_{k,M}$ and $v'_{k,M}$ have to be replaced by $v_{k,P}$ and $v'_{k,P}$, where the M are replaced by P.

Proof

We look at the possible frequencies of words from $n = o_{k-1} + 1$ to $n = o_k$; suppose $\eta_k = -1$. Then $r_{1,n}$ is Δ_k from $n = o_{k-1} + 1$ to $n = o_{k-1} + n_k$, then $r_{1,n}$ is r_{1,o_k} from $n = o_{k-1} + n_k + 1$ to $n = o_k$; $r_{2,n}$ is Δ_k from $n = o_{k-1} + 1$ to $n = o_{k-1} + m_k$, then $r_{2,n}$ is r_{2,o_k} from $n = o_{k-1} + m_k$ to $n = o_k$; $l_{1,n}$ takes values which are (by definition of the induction) larger than $r_{1,n}$ from $n = o_{k-1} + 1$ to $n = o_{k-1} + n_k - 1$, then $l_{1,n}$ is l_{1,o_k} from $n = o_{k-1} + n_k$ to $n = o_k$. $l_{2,n}$ takes values which are (by definition of the induction) larger than $r_{2,n}$ from $n = o_{k-1} + 1$ to $n = o_{k-1} + m_k - 1$, then $l_{2,n}$ is l_{2,o_k} from $n = o_{k-1} + m_k$ to $n = o_k$.

By Lemmas 3.5 and 3.6, we get that the smallest value of ne_n between $n = |w_{1,o_{k-1}}| + 2$ and $n = |w_{1,o_k} + 1|$ is reached by one of five quantities:

- $pr_{1,n} = pr_{2,n}$ for $p = |w_{1,o_{k-1}}| + 2$, $n = o_{k-1} + 1$,
- $pl_{1,n}$ for $p = |w_{1,o_{k-1}+n_k-1}| + 2$, $n = o_k$,
- $pl_{2,n}$ for $p = |w_{2,o_{k-1}+m_k-1}| + 2$, $n = o_k$,
- $pr_{1,n}$ for $p = |w_{1,o_{k-1}+n_k}| + 2$, $n = o_k$,
- $pr_{2,n}$ for $p = |w_{2,o_{k-1}+m_k}| + 2$, $n = o_k$.

Note that if $n_k = 1$, the first quantity is smaller than the second, and could be dispensed with, but it is easier to use all the five quantities than to make many special subcases.

The first one is $z_{k-1}\Delta_k$, which gives formula (14) in view of (11).

The other ones are $l_{1,o_k}|w_{1,o_{k-1}+n_k-1}|$, $l_{2,o_k}|w_{2,o_{k-1}+m_k-1}|$, $r_{1,o_k}|w_{1,o_{k-1}+n_k}|$, $r_{2,o_k}|w_{2,o_{k-1}+m_k}|$.

We deal first with the third and fourth ones: we write (3) for $n = o_k - 1$; we get that $(l_{1,o_k} + l_{2,o_k})z_{k-1} + \Delta_k(|w_{1,o_{k-1}+n_k}| + |w_{2,o_{k-1}+n_k}| - 2z_{k-1})$, which is equal to $\Delta_k(|w_{1,o_{k-1}+n_k}| + |w_{2,o_{k-1}+m_k}|) - (r_{1,o_k} + r_{2,o_k})z_{k-1}$, is close to 1; we write

$$r_{1,o_k}|w_{1,o_{k-1}+n_k}| = (r_{1,o_k}+r_{2,o_k})(|w_{1,o_{k-1}+n_k}|+|w_{2,o_{k-1}+m_k}|)\frac{r_{1,o_k} + r_{2,o_k}}{r_{1,o_k}}\frac{|w_{1,o_{k-1}+n_k}| + |w_{2,o_{k-1}+m_k}|}{|w_{1,o_{k-1}+n_k}|},$$

and notice that $|w_{1,o_{k-1}+n_k}| + |w_{2,o_{k-1}+m_k}|$ is close to $z_k + z_{k-1}$; after deducing $\frac{\Delta_k}{r_{1,o_k}+r_{2,o_k}}$ from (6) and imputting the values of the lengths of the bispecial words, this gives (17), and (18) is similar.

The expression (3) for $n = o_k - 1$ is also close to $(l_{1,o_k} + l_{2,o_k})z_{k-1} + \Delta_k(|w_{1,o_{k-1}+n_k-1}| + |w_{2,o_{k-1}+m_k-1}|)$, which is thus close to 1; as $|w_{1,o_{k-1}+n_k-1}|+|w_{2,o_{k-1}+m_k-1}|$ is close to $z_k - z_{k-1}$, this gives (15) and (16).

And the case $\eta_k = 1$ is deduced as usual. □

To some extent, the quantities B_k and B'_k measure the quality of approximation of the angle of the inducing rotation by the semi-regular continued fraction (8), which is best if t_k or v_k is close to 0 or 1, see Proposition 5.1 of [15].

The value of $\mathcal{B}$ is not changed if we change a finite number of parameters of the expansion; this will be understated in the sequel when not recalled explicitely.

We remark that v_k depends only on $m_j + n_j$ for $j \le k$ and $-\eta_j\eta_{j+1}$ for $j \le k-1$; $\eta_k\eta_{k+1}t_k$ depends only on $m_j + n_j$ for $j \ge k+1$ and $-\eta_j\eta_{j+1}$ for $j \ge k$; $t_{k,l}$ and $t_{k,r}$ depend only on (n_j, m_j) for $j \ge k+1$ and $-\eta_j\eta_{j+1}$ for $j \ge k$; the $v_{k,.}$ and $v'_{k,.}$ depend only on (n_j, m_j) for $j \le k$ and $-\eta_j\eta_{j+1}$ for $j \le k-1$.

Proposition 3.8. *$\mathcal{B}'$ is the smallest of the lower limits, when $k \to \infty$, of the following quantities*

- *when $\eta_k = -$ and $|w_{2,o_{k-1}+m_k-1}| \le |w_{2,o_{k-1}+m_k}| \le |w_{1,o_{k-1}+n_k-1}| \le |w_{1,o_{k-1}+n_k}|$:* $\frac{1}{\Delta_k|w_{2,o_{k-1}+m_k-1}|}$, $\frac{1}{l_{2,o_k}|w_{2,o_{k-1}+m_k}|}$, $\frac{1}{\min(l_{2,o_k},r_{2,o_k})|w_{1,o_{k-1}+n_k-1}|}$, $\frac{1}{\min(l_{1,o_k},l_{2,o_k},r_{2,o_k})|w_{1,o_{k-1}+n_k}|}$, $\frac{1}{\min(l_{1,o_k},l_{2,o_k},r_{1,o_k},r_{2,o_k})z_k}$;
- *when $\eta_k = -$ and $|w_{2,o_{k-1}+m_k-1}| \le |w_{1,o_{k-1}+n_k-1}| \le |w_{2,o_{k-1}+m_k}| \le |w_{1,o_{k-1}+n_k}|$:* $\frac{1}{\Delta_k|w_{2,o_{k-1}+m_k-1}|}$, $\frac{1}{l_{2,o_k}|w_{1,o_{k-1}+n_k-1}|}$, $\frac{1}{\min(l_{1,o_k},l_{2,o_k})|w_{2,o_{k-1}+m_k}|}$, $\frac{1}{\min(l_{1,o_k},l_{2,o_k},r_{2,o_k})|w_{1,o_{k-1}+n_k}|}$, $\frac{1}{\min(l_{1,o_k},l_{2,o_k},r_{1,o_k},r_{2,o_k})z_k}$;
- *as in the first two cases with 1 and 2 exchanged, m and n exchanged;*
- *as in the first four cases with r and l exchanged and $\eta_k = +$.*

Proof
It is a straightforward consequence of the proof of Proposition 3.7. □

We show how to compute effectively $\mathcal{B}$ for a particular family of cases:

Proposition 3.9. *Let $n \geq m \geq 1$ be two integers; we consider a three-interval exchange such that for all k $n_k = n$, $m_k = m$, $\eta_k\eta_{k+1} = -1$. Then t is the smaller root of the polynomial $X^2 + (m+n)X - 1$, and $\mathcal{B}$ is the largest of the three quantities*

$$t + \frac{1}{t}, \quad \left(\frac{(m+n)^2}{m^2}\right)\left(\frac{1}{1+t} + \frac{t}{1-t}\right),$$
$$\left(\frac{2(1-t)^2}{(1-t)^2 + (m-n)t}\right)\left(\frac{2(1+t)^2}{(1+t)^2 + (n-m)t}\right)\left(\frac{1}{1+t} + \frac{t}{1-t}\right).$$

Proof
Then for every k we have $t_k = v_k = t$. Thus $r_{1,o_k} + r_{2,o_k} = \Delta_k(1 - \eta_{k+1})t$, while $r_{1,o_k} - r_{2,o_k} = \Delta_k\eta_{k+1}u$, for some constant u; by (5) we get $u = \frac{(n-m)t}{1+t}$. Similarly, up to higher order terms, $z_{k-1} = tz_k$, and $|M_{1,o_k}| - |M_{2,o_k}| = z_k u'$ with $u' = \frac{(n-m)t}{1-t}$.

Without loss of generality, we suppose $\eta_k = -1$, and compute the various quantities in formulas (14) to (18). Formula (14) becomes $t + \frac{1}{t}$, and $B_k = B'_k = \frac{1}{1+t} + \frac{t}{1-t}$. Note also that formula (17) gives a smaller estimate than (18) and thus will not be used.

We look now at $1 + \frac{1}{t_{k,r}}$; inputting the above values, we get that it is equal to $\frac{2-2t^2}{1-t^2+(m-n)t}$; but, using the equation defining t, we check that it is just equal to $\frac{m+n}{m}$. Similarly, we get $1 + t_{k,l} = \frac{2(1+t)^2}{(1+t)^2+(m-n)t}$ and $1 + \frac{1}{t_{k,l}} = \frac{2(1+t)^2}{(1+t)^2+(n-m)t}$, and these do not seem to admit simpler expressions.

In the same way, we get that $1 + \frac{1}{v'_{k,M}}$ is equal to $\frac{2+\frac{2}{t}}{2m+1+t+(m-n)\frac{t}{1-t}}$; this last quantity turns out to be also equal to $\frac{m+n}{m}$. Then $1 + v_{k,M}$ is equal to $\frac{-2+\frac{2}{t}}{2n-1+t+(n-m)\frac{t}{1-t}}$, which, by using the equation defining t, is also equal to $\frac{2(1-t)^2}{(1-t)^2+(n-m)t}$. Finally $1 + \frac{1}{v_{k,M}}$ can be recovered from the last quantity and is equal to $\frac{2(1-t)^2}{(1-t)^2+(m-n)t}$. But at this stage we notice that formula (15) gives a smaller estimate than (16) and thus will not be used. Thus we have the three claimed formulas. □

3.4. **The upper BL spectrum: smallest elements.**

Theorem 3.10. *The smallest element in the upper BL spectrum of three-interval exchange transformations, and the only one below $\frac{12+29\sqrt{3}}{13} = 4,786...$, is $2\sqrt{5} = 4,47....$ It is reached if and only if for all k large enough $m_k = n_k = 2$ and $\eta_k\eta_{k+1} = -$.*
The spectrum is a closed set and contains an accumulation point equal to 6.

Proof
We look at formulas (14) to (18). The maximum of (15) and (16) is at least $4B_k$, while the maximum of (17) and (18) is at least $4B'_k$.

We take a three-interval exchange with expansion (n_k, m_k, η_k) and look at the sequence (a_k, ε_k) where $a_k = m_k + n_k$, $\varepsilon_k = -\eta_k\eta_{k+1}$.

If there are infinitely many $a_k \geq 6$, or $(a_k, \varepsilon_k) = (5, +)$, then for infinitely many k formula (14) gives a bound greater or equal to 5.
If there are only $a_k \leq 4$ or $(a_k, \varepsilon_k) = (5, -)$, and for infinitely many k we have $\varepsilon_k = -$; then we have always $t_k \geq [0, (5 * -)^\omega] = \frac{5-\sqrt{21}}{2}$ and also $v_k \geq [0, (5 * -)^\omega]$. Then if $\varepsilon_k = -$, $4B_k$ is at least $\frac{4\sqrt{21}}{3} = 6, 11...$
If there are only $(a_k, \varepsilon_k) = (3, +)$, $(a_k, \varepsilon_k) = (4, +)$ or $(a_k, \varepsilon_k) = (2, +)$, with infinitely many $(2, +)$; then we have always $v_k \geq [0, (2 * +, 4 * +)^\omega] = \sqrt{6} - 2$. If $(a_{k+1}, e_{k+1}) = (2, +)$, then $t_k \geq [0, 2 * +(2 * +, 4 * +)^\omega] = \frac{1}{\sqrt{6}}$ and thus $4B'_k$ is at least $\frac{8+8\sqrt{6}}{5} = 5, 519....$
If there are only $(a_k, \varepsilon_k) = (3, +)$ or $(a_k, \varepsilon_k) = (4, +)$, with infinitely many $(3, +)$; then we have always $v_k \geq [0, (3 * +, 4 * +)^\omega] = \frac{4\sqrt{3}-6}{3}$, If $(a_{k+1}, e_{k+1}) = (3, +)$, then $t_k \geq [0, 3 * +(3 * +, 4 * +)^\omega] = \frac{4\sqrt{3}-3}{13}$ and thus $4B'_k$ is at least $\frac{12+29\sqrt{3}}{13} = 4, 786....$
If there are only $(a_k, \varepsilon_k) = (4, +)$, with for infinitely many k $m_k \neq n_k$; then $t_k = v_k = \sqrt{5} - 2$. We take a k such that $m_k \neq n_k$; without loss of generality we take $\eta_k = -$. If $m_k = 1$ and $n_k = 3$, then $r_{1,o_k} > r_{2,o_k}$, and $\frac{1}{v'_{k,M}} = \frac{3z_{k-1}+|M_{1,o_{k-1}}|}{z_{k-1}+|M_{2,o_{k-1}}|} \geq \frac{3z_{k-1}}{z_{k-1}+|M_{1,o_{k-1}}|+|M_{2,o_{k-1}}|} = \frac{3z_{k-1}}{2z_{k-1}+z_{k-2}}$; as $z_{k-2} = v_{k-1}z_{k-1}$, formula (18) gives at least $\frac{15+7\sqrt{5}}{4} = 7, 663...$ If $m_k = 3$ and $n_k = 1$, formula (17) gives the same estimate.

There remains only the case $(a_k, \varepsilon_k) = (4, +)$ and $m_k = n_k$, where $\mathcal{B}$ has the claimed value; note that each of the formulas (14) to (18) gives the same result.

If we take (a_k, ε_k) to be $((2, +)^2, (4, +)^j)^\omega$, and $m_k = n_k$, the value of $\mathcal{B}$ is reached (among others) by $4B'_k$ for $a_k = 2$ and $a_{k-1} = 4$; when j tends to infinity, these values of v_k and t_k both tend to $[0, 2 * +(4 * +)^\omega] = \frac{1}{\sqrt{5}}$, and thus $\mathcal{B}$ tends to 6. The spectrum is closed by the standard reasoning of [14], Chapter 1, Corollary to Lemma 6, as every finite value is reached with bounded m_k and n_k, and for any sequence of three-interval exchanges such that the corresponding $\mathcal{B}$ converge to a finite number, we can take the m_k and n_k uniformly bounded (and $+\infty$ is in the spectrum, see Proposition 3.11 below). 6 is indeed an accumulation point, approached from below if j is even. $\square$

The above computations illustrate how $\mathcal{B}$ is easiest to compute when $m_k = n_k$ for all k; when $m_k \neq n_k$, the situation becomes much more complicated; indeed, some of the bounds in the above computations may look quite crude, as for example if $m_k + n_k = 3$ we cannot have $m_k = n_k$ and the bounds $4B_k$ and $4B'_k$ are never reached, but in general they are not easy to improve.

In view of the considerations above, the values of $\mathcal{B}$ between $2\sqrt{5}$ and 6 were found only by trial and error; the second value is very likely to be $4\sqrt{2} = 5, 65...$, which is reached when $(n_k, m_k, -\eta_k\eta_{k+1})$ is either $(1, 1, +)^\omega$ or $(3, 3, -)^\omega$; the latter gives the minimal value for m_k and n_k constant and $\eta_k\eta_{k+1} = +1$. The third value we found is $\frac{20\sqrt{26}}{17} = 5, 9988...$, which is reached for $((1, 1, -)^2, (2, 2, -)^2)^\omega$ and the fourth one is $\frac{4\sqrt{209306}}{305} = 5, 999996...$, which is reached for $((1, 1, -)^2, (2, 2, -)^4)^\omega$.

Thus the first, second, third, fourth smallest element we found in the upper BL spectrum of three-interval exchange transformations is respectively twice the first, second, sixth and twelfth Lagrange number. Though of course we might have missed some values, it seems likely that *the upper BL spectrum of three-interval exchanges below* 6 *is strictly included in* *twice the Lagrange spectrum below* 3; thus we conjecture that 6 *is the lowest accumulation point of our spectrum.*

When $(n_k, m_k, -\eta_k\eta_{k+1})$ is $(2,2,+)^\omega$, the values of $\alpha = \beta$ and the angle of the inducing rotation can be deduced from (8); we get $\alpha = \beta = \frac{3-\sqrt{5}}{2}$, and the inducing rotation is of angle $\frac{1}{\sqrt{5}}$; as in that case (8) gives a classical continued fraction expansion, the $\mathcal{B}$ of the rotation is computed by Theorem 2.4, and we check that the $\mathcal{B}$ of the interval exchange is the one of the inducing rotation. By similar computations, for $(1,1,+)^\omega$, the $\mathcal{B}$ of the interval exchange is twice the one of the inducing rotation. For $((1,1,+)^2,(2,2,+)^j)^\omega$ with j large, the $\mathcal{B}$ of the interval exchange is close to 6, while the one of the inducing rotation is between 4 and 5.

A partial analogy with the Lagrange spectrum is that in the latter the sequence increasing to 3 is given by angles whose Euclid continued fraction expansion is $[0,(2^2 1^{2j})^\omega]$, while in our upper BL spectrum the sequence increasing to 6 is given by quantities t_k whose Euclid continued fraction expansion is $[0,(2^2 4^{2j})^\omega]$; but it fails for the angles whose classic continued fraction expansion is $[0,(2^4 1^2)^\omega]$, which gives a rotation with $\mathcal{B} = 2,9992....$, while the three-interval exchange for which $(n_k, m_k, -\eta_k\eta_{k+1})$ is $((1,1,+)^4,(2,2,+)^2)^\omega$ has $\mathcal{B} = 6,06.....$

All the values of $\mathcal{B}$ we have found below 6 are reached by the estimates $4B_k$ or $4B'_k$, though when $(n_k, m_k, -\eta_k\eta_{k+1})$ is either $(2,2,+)^\omega$ or $(3,3,-)^\omega$ they are also reached by formula (14). The smallest value we found without $m_k = n_k$ for all k large enough is $\frac{5\sqrt{29}}{4} = 6,73...$, which is reached when $(n_k, m_k, -\eta_k\eta_{k+1})$ is $(3,2,+)^\omega$; for $(3,1,+)^\omega$ we get $\mathcal{B} = 8\sqrt{5}$.

3.5. **The upper BL spectrum: largest elements.**

Proposition 3.11. *For a three-interval exchange, $\mathcal{B} = +\infty$ if and only if the angle of the inducing rotation has unbounded partial quotients for the Euclid algorithm.*

Proof

By the formulas in Proposition 3.7, $\mathcal{B} = +\infty$ if and only if either $m_k + n_k$ is unbounded, or t_k or v_k takes values arbitrarily close to one. The last two possibilities are equivalent to the existence of unbounded strings of $m_k + n_k = 2$, $-\eta_k\eta_{k+1} = -$, and this gives the result by (8) and Section 1.3. □

The upper BL spectrum above 6 seems to become quite complicated, maybe more than the Lagrange spectrum above 3. However, we can find enough values given by formula (14) (this happens when all $m_k = n_k$ and some of them are large enough) to fill an interval $[C, +\infty[$. This involves adapting to semi-regular continued fraction expansions the famous Theorem 3.1 of [23]:

Lemma 3.12. *Let G be the set of numbers $[0, a_1 * \varepsilon_1, ..., a_n * \varepsilon_n, ...]$ where all the a_i are 2, 4, 6 or 8, the ε_i are $-$ or $+$, and there are never two consecutive $2*-$; then $G+G$ is the interval $[\frac{2\sqrt{195}-24}{17}, \frac{2\sqrt{195}-24}{3}] = [0,23108..., 1,30949...]$, and $G-G$ is the interval $[\frac{168-14\sqrt{195}}{51}, \frac{14\sqrt{195}-168}{51}] = [-0,53920..., 0,53920...]$.*

Proof

We follow the proof of Theorem 2 in Chapter 4 of [14], which in turns follows [23]. The largest number in G is $[0,(2*-,2*+,8*+)^\omega] = \frac{\sqrt{195}-12}{3}$, and the smallest one is $[0,(8*+,2*-,2*+)^\omega] = \frac{\sqrt{195}-12}{17}$. We can build G as a Cantor set in the following way: we start form the interval $[\frac{\sqrt{195}-12}{17}, \frac{\sqrt{195}-12}{3}]$. Then, for each sequence $(b_1, \varepsilon_1), ...(b_n, \varepsilon_n)$ with no two consecutive $(2, -)$ and $b_i = 2, 4, 6, 8$: if the number of ε_i which are $+$ is 0 or even, we delete the intervals $][0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, a * +, (8 * +, 2 * -, 2 * +)^\omega], [0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, a * -, (8 * +, 2 * -, 2 * +)^\omega][$ for $a = 2, 4, 6, 8$, and $][0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, (a + 2) * -, (2 * -, 2 * +, 8 * +)^\omega], [0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, a * +, (2 * -, 2 * +, 8 * +)^\omega][$ for $a = 2, 4, 6$; if the number of ε_i which are $+$ is odd, we delete the intervals $][0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, a * -, (8 * +, 2 * -, 2 * +)^\omega], [0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, a * +, (8 * +, 2 * -, 2 * +)^\omega][$ for $a = 2, 4, 6, 8$, and $][0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, a * +, (2 * -, 2 * +, 8 * +)^\omega], [0, b_1 * \varepsilon_1, ..., b_n * \varepsilon_n, (a + 2) * -, (2 * -, 2 * +, 8 * +)^\omega][$ for $a = 2, 4, 6$. We check that this dissection can be done in successive stages, where each time we delete the interval A_2 from the interval $A_1 \cup A_2 \cup A_3$, and the length of A_2 is smaller than the length of A_1 and than the length of A_3; then Lemmas 2 to 4 of [14] give the result. The same reasoning works for $G+(-G)$. $\square$

Theorem 3.13. *The upper BL spectrum of the family of three-interval exchanges contains the interval* $[12 + 2\sqrt{2} = 14, 828..., +\infty]$.

Proof

By Proposition 3.11, the upper BL spectrum contains $+\infty$. Let s be a real number larger than $12 + 2\sqrt{2}$; by Lemma 3.12, we can write $s = r + \varepsilon[0, a_1 * \varepsilon_1, ..., a_n * \varepsilon_n, ...] + \varepsilon'[0, a'_1 * \varepsilon'_1, ..., a'_n * \varepsilon'_n, ...]$, where r is an even positive integer, all the a_i and a'_i are 2, 4, 6 or 8, the ε_i, ε'_i, ε and ε' are $-$ or $+$, and there are never two consecutive $2 * -$ in the $a_n * \varepsilon_n$ and $a'_n * \varepsilon'_n$. We choose some increasing sequence k_n such that $a'_{k_n} * \varepsilon'_{k_n}$ is not $2 * -$.

We define the expansion $(m_k + n_k, -\eta_k\eta_{k+1})$ to be $(a_{k_1}, \varepsilon_{k_1-1}), ...(a_3, \varepsilon_2)(a_2, \varepsilon_1)(a_1, \varepsilon)(r, \varepsilon')$ $(a'_1, \varepsilon'_1)...(a'_{k_1}, \varepsilon'_{k_1})(a_{k_2}, \varepsilon_{k_2-1}), ...(a_3, \varepsilon_2)(a_2, \varepsilon_1)(a_1, \varepsilon)(r, \varepsilon')(a'_1, \varepsilon'_1)...(a'_{k_2}, \varepsilon'_{k_1})...$ We take now the three-interval exchange for which $m_k + n_k$ and $-\eta_k\eta_{k+1}$ are defined in that way, and $m_k = n_k$ for all k; it exists as this gives an admissible expansion. By Proposition 3.7 $\mathcal{B}$ is the maximum of the upper limits of $\frac{1}{v_k} - \eta_k\eta_{k+1}t_k$, and of $4B_k$ and $4B'_k$. As r is at least 14, the first of these upper limits can be taken on those k for which $m_k + n_k = r$, and is exactly s.

We look now at the other two upper limits, for which we need an upper bound for t_k and v_k. If we compute t_k when $(m_k + n_k, -\eta_k\eta_{k+1})$ is inside a string of (a'_n, ε'_n), we do not see two consecutive $(2, -)$ in that string, while, when we start inside a string of (a_n, ε_n), in that string we do not see $(a, -)$ followed by $(2, -)$ followed by $(2, f)$ for any a and f, and the two cases are exchanged if we compute v_k; in both cases starting from a (r, ε') or near a junction of two different strings does not change the bounds. Thus either t_k is at most $[0, 2 * -, 2] = \frac{2}{3}$, and v_k is at most $[0, 2 * -, (2 * -, 4 * -)^\omega] = \frac{1}{\sqrt{2}}$, or v_k is at most $[0, 2 * -, 2] = \frac{2}{3}$, and t_k is at most $[0, 2 * -, (2 * -, 4 * -)^\omega] = \frac{1}{\sqrt{2}}$; these bounds can be improved when $(m_k + n_k, -\eta_k\eta_{k+1}) = (2, -)$, then either t_k is at most $[0, 2] = \frac{1}{2}$, and v_k is at most $[0, 2 * -, (4 * -, 2 * -)^\omega] = 2 - \sqrt{2}$, or v_k is at most $[0, 2 * -, 2] = \frac{2}{3}$, and t_k is at most $[0, (2 * -, 4 * -)^\omega] = 2 - \sqrt{2}$; also, when $m_k + n_k$ is at least 4, v_k is at most either $[0, 4 * -, (2 * -, 4 * -)^\omega] = \frac{2-\sqrt{2}}{2}$, or $[0, 4 * -, 2 * -, 2] = \frac{3}{10}$. We input these values in the quantities we want to bound, together with $t_k > 0$, $v_k > 0$; the worst case happens when $(m_k + n_k, -\eta_k\eta_{k+1}) = (2, -)$, and we get that the last two upper limits are not larger than $12 + 2\sqrt{2} \le s$. $\square$

The lower bound on the interval we give should certainly be improved, but let us point that this would need a different method, as, to improve significantly the bound in Lemma

3.12, we would need to replace G by a set where the expansions do not contain $2*-$, and such a set does not contain elements below $\frac{1}{2}$.

3.6. **The lower BL spectrum.** Proposition 3.8 allows us to compute $\mathcal{B}'$ for individual interval exchanges, by the usual trick of replacing the 1 in the numerators by the quantity in relation (11), thus we have to estimate ratios of lengths and frequencies; for example, for $m_k = n_k = 2$ and $\eta_k\eta_{k+1} = -$, we have $l_{1,o_k} = l_{2,o_k}$, $r_{1,o_k} = r_{2,o_k}$; $\mathcal{B}'$ is the smallest of the lower limits (taken for $\eta_k = -$) of $\frac{1}{\Delta_k|w_{2,o_{k-1}+m_k-1}|}$, $\frac{1}{l_{2,o_k}|w_{2,o_{k-1}+m_k}|}$, $\frac{1}{r_{2,o_k}z_k}$; taking into account that $|M_{2,o_{k-1}}| = |M_{1,o_{k-1}}|$, we check that these three quantitites are all equal and that $\mathcal{B}' = 5 - \sqrt{5} = 2,763...$

But if we want to compute $\mathcal{B}'$ for a general three-interval exchange, these formulas are complicated and do not reduce to simpler expressions as for $\mathcal{B}$. However, the lower BL spectrum for the family is fully known, and quite different from the lower BL spectrum of rotations.

Theorem 3.14. *The lower BL spectrum of the family of three-interval exchanges is the interval* $[2, +\infty]$.

Proof

No number smaller than 2 can be in the lower BL spectrum because of Lemma 1.2. We take $2 \leq s \leq +\infty$, and choose integers $m_k < n_k$ such that m_k and n_k grow to infinity with k, and $\frac{m_k+n_k}{m_k} \to s$ when k goes to infinity; we require $n_k - m_k \to +\infty$, which is an additional condition when $s = 2$; we choose $\eta_k = -$ for all k.

Then by Lemmas 3.3 and 3.4, we have $|w_{2,o_{k-1}+m_k-1}| \leq |w_{2,o_{k-1}+m_k}| \leq |w_{1,o_{k-1}+n_k-1}| \leq |w_{1,o_{k-1}+n_k}|$, $l_{2,o_k} \leq r_{1,o_k} \leq r_{2,o_k}$, $l_{2,o_k} \leq l_{1,o_k} \leq r_{2,o_k}$. By Proposition 3.8, $\mathcal{B}'$ is the smaller of the lower limits of $\frac{1}{\Delta_k|w_{2,o_{k-1}+m_k-1}|}$ and $\frac{1}{l_{2,o_k}z_k}$.

Now $|w_{2,o_{k-1}+m_k-1}|$ is equivalent to $m_k z_{k-1}$ while, by the computations in the proof of Proposition 3.7 $\frac{1}{\Delta_k z_{k-1}}$ is close to v_k (because $t_k \to 0$) which is equivalent to $m_k + n_k$, thus the first lower limit is a limit and is equal to s. The same is true for the second one as $\frac{1}{\Delta_{k+1}z_k}$ and $\frac{\Delta_k}{\Delta_{k+1}}$ are equivalent to $m_{k+1} + n_{k+1}$, and, because $r_{1,o_k} + r_{2,o_k}$ is equivalent to Δ_k and $r_{1,o_k} - r_{2,o_k}$ is equivalent to $(m_{k+1} - n_{k+1})\Delta_{k+1}$, we get that $\frac{l_{2,o_k}}{\Delta_k}$ is equivalent to $\frac{m_{k+1}}{m_{k+1}+n_{k+1}}$. $\square$

Thus for some uniquely ergodic three-interval exchange transformations we have $ne_n \to 0$ when n tends to infinity; this result, and its consequence that Boshernitzan's criterion is not a necessary condition in this family of systems, are stated without proof in [27]. Note that the covering number by intervals (see the opening of Section 2 above) of a three-interval exchange is shown in [5] to be the same as for the inducing rotation, and thus is not equal to $\frac{1}{\mathcal{B}'}$, in contrast with the case of rotations.

4. Arnoux-Rauzy systems

The *Arnoux-Rauzy systems* are defined in [3] as the minimal symbolic systems on the alphabet $\{1, 2, 3\}$ such that the complexity of the language is $2n + 1$ for all n, and, for all n, there are one right special and one left special word. Then [3] proceeds to give a constructive (additive) algorithm to generate them with three families of words, built with three rules denoted by a, b and c; [11] gives a multiplicative version of this construction, which we take here as a definition, valid up to permutations of $\{1,2,3\}$: the k_n are the number of consecutive times a given rule is used, while the $n_i > 1$ mark the times where three consecutive rules are all different, such as, up to permutations of $\{a,b,c\}$, rule a used k_{n_i-1} times, then rule b used k_{n_i} times, then rule c used k_{n_i+1} times.

Definition 4.1. *Given two infinite sequences of integers $k_n \geq 1$, $n \geq 1$, and $n_1 < n_2 ... < n_i < ...$ the Arnoux-Rauzy system (X_L, S) defined by them is the symbolic system associated to the language L of all factors of $(H_n)_{n\in\mathbb{N}}$, where the three words H_n, G_n, J_n are built from $H_0 = 1$, $G_0 = 2$, $J_0 = 3$ by two families of rules:*

- *if $n+1 = n_i$ for some i, $H_{n+1} = G_n H_n^{k_{n+1}}$, $G_{n+1} = J_n H_n^{k_{n+1}}$, $J_{n+1} = H_n$;*
- *otherwise, $H_{n+1} = G_n H_n^{k_{n+1}}$, $G_{n+1} = H_n$, $J_{n+1} = J_n H_n^{k_{n+1}}$.*

Every Arnoux-Rauzy system is minimal [3] and uniquely ergodic (by [7] because the complexity is $2n+1$). Though they are defined as symbolic systems, they have also geometric models, see [1] [2] [3][25]: every Arnoux-Rauzy system is a coding of a six-interval exchange on the circle, and some of them are codings of rotations of the 2-torus.

Proposition 4.1. *Let $o_n = k_1 + ... k_n$, $n > 0$, and $o_0 = 0$. All the bispecial words of the language are the w_n, where w_0 is the empty word and for $1 \leq j \leq k_{n+1}$, $w_{o_n+j} = w_{o_n} H_n^j$. For $1 \leq j \leq k_{n+1}$, let $\theta_{n,j}$, resp. $\gamma_{n,j}$, $\iota_{n,j}$, be the frequency of $w_{o_n+j}x$ where x is the first letter of H_{n+1}, resp. G_{n+1}, J_{n+1}; we denote $\theta_{n,k_{n+1}}$, resp. $\gamma_{n,k_{n+1}}$, $\iota_{n,k_{n+1}}$, by θ_{n+1}, resp. γ_{n+1}, ι_{n+1}, then*

- *if $n+1 = n_i$ for some i, $\theta_n = j\theta_{n,j} + j\gamma_{n,j} + \iota_{n,j}$, $\gamma_n = \theta_{n,j}$, $\iota_n = \gamma_{n,j}$,*
- *otherwise $\theta_n = j\theta_{n,j} + \gamma_{n,j} + j\iota_{n,j}$, $\gamma_n = \theta_{n,j}$, $\iota_n = \iota_{n,j}$.*

Proof
This is a straightforward consequence of [3], taking into account the multiplicative definition, as in Proposition 9 of [11]. □

We check that we have the relations

$$|H_n|\theta_n + |G_n|\gamma_n + |J_n|\iota_n = 1, \tag{19}$$

$$|H_n| + |G_n| + |J_n| - 2|w_{o_n}| = 3. \tag{20}$$

We shall use also Lemma 7 of [11], with the more precise (and non-trivial) estimate used in its proof: namely, for all n, if $(t_{1,n}, t_{2,n}, t_{3,n})$ is the triple $(|J_n|, |G_n|, |H_n|)$ ordered so that $t_{1,n} \leq t_{2,n} \leq t_{3,n}$, then we have $t_{2,n} \leq t_{1,n} + t_{3,n}$.

Proposition 4.2. *For an Arnoux-Rauzy system, if $S_n = |H_n| + |G_n| - |J_n|$,*

$$\mathcal{B} = \limsup_{i\to+\infty} \left(\frac{2|H_{n_i}|}{S_{n_i}} \frac{\theta_{n_i}}{\theta_{n_{i+1}+1}} + \frac{2|G_{n_i}|}{S_{n_i}} \frac{\theta_{n_i+1}}{\theta_{n_{i+1}+1}} + \frac{2|J_{n_i}|}{S_{n_i}} \right),$$

$$\mathcal{B}' = \liminf_{i\to+\infty} \left(\frac{2|H_{n_i}|}{S_{n_{i+1}}} \frac{\theta_{n_i}}{\theta_{n_{i+1}+1}} + \frac{2|G_{n_i}|}{S_{n_{i+1}}} \frac{\theta_{n_i+1}}{\theta_{n_{i+1}+1}} + \frac{2|J_{n_i}|}{S_{n_{i+1}}} \right).$$

Proof
On the Rauzy graph of length m, there are at most four frequencies, which are for example those of $w_{o_n+j}1$, $w_{o_n+j}2$, $w_{o_n+j}3$, and w_{o_n+j}, thus $\theta_{n,j}$, $\gamma_{n,j}$, $\iota_{n,j}$ and $\theta_{n,j} + \gamma_{n,j} + \iota_{n,j}$, for the

smallest n and j such that $|w_{o_n+j}| \geq m$. In the same way as what happens in Lemma 2.3, for $m = |w_p| + 1$ there are only three frequencies, which are equal to the three lowest frequencies for $m = |w_p|$, and for $m \neq |w_p| + 1$ all the four frequencies appear.

For all n, the formulas imply that $\theta_n \geq \gamma_n$ and $\theta_n \geq \iota_n$ We have always $\theta_{n+1} = \gamma_n$ and either $\gamma_{n+1} = \iota_n$ or $\iota_{n+1} = \iota_n$, thus always $\gamma_n \geq \iota_n$. Then for $1 \leq j \leq k_{n+1} - 1$, we have either $\theta_{n,j} = \theta_{n+1}$, $\gamma_{n,j} = \gamma_{n+1} = \iota_n$, $\iota_{n,j} = \iota_{n+1} + (k_{n+1} - j)(\theta_{n+1} + \gamma_{n+1})$, or $\theta_{n,j} = \theta_{n+1}$, $\iota_{n,j} = \iota_{n+1} = \iota_n$, $\gamma_{n,j} = \gamma_{n+1} + (k_{n+1} - j)(\theta_{n+1} + \iota_{n+1})$. Hence for all $0 \leq j \leq k_{n+1} - 1$, $\min(\theta_{n,j}, \gamma_{n,j}, \iota_{n,j}) = \iota_n$.

Taking into account that $\iota_{n+1} = \iota_n$ if and only if $n + 1$ is not an n_i, we get that for every $i \geq 1$, e_r is equal to ι_{n_i} from $r = |w_{o_{n_i}-1}| + 2$ to $r = |w_{o_{n_{i+1}}-1}| + 1$.

We have $|w_{o_{n_i}-1}| = |w_{o_{n_i}}| - |H_{n_i-1}| = |w_{o_{n_i}}| - |J_{n_i}|$, and this is $\frac{|H_{n_i}|+|G_{n_i}|-|J_{n_i}|}{2}$ after discarding the constant term, which gives $\mathcal{B} = \limsup_{i\to+\infty} \frac{2}{\iota_{n_i} S_{n_i}}$, and $\mathcal{B}' = \liminf_{i\to+\infty} \frac{2}{\iota_{n_i} S_{n_{i+1}}}$. We now input the relation (19), and the fact that for any m, $\gamma_m = \theta_{m+1}$ and, if m' is the smallest $n_i > m$, $\iota_m = \iota_{m'-1} = \gamma_{m'} = \theta_{m'+1}$, to get the final formulas. □

Note that in the formula giving $\mathcal{B}$, the ratios of lengths of words depend only on the n_j for $j \leq i$ and the k_t for $t \leq n_i$, while the ratios of frequencies depend only on the n_j for $j > i$ and the k_t for $t > n_i$; in the formula giving $\mathcal{B}'$, there is no such dichotomy. To every Arnoux-Rauzy system is associated an algoritm of simultaneous approximation of two irrationals, see [3] [11]; this involves the θ_n and the $|H_n|$, and the approximation is best when the k_n are large; this algorithm is hidden in all the computations of this section, though only $\mathcal{B}$ is linked to the quality of the approximation, and only in a loose way.

Proposition 4.3. *The upper BL spectrum of the family of Arnoux-Rauzy systems contains* $+\infty$*, which is reached if and only if the* k_n*,* $n \in \mathbb{N}$*, or the* $n_{i+1} - n_i$*,* $i \geq 1$*, are unbounded. Its smallest element, and the only one below* $\frac{181}{21} = 8,619...$*, is reached for the* Tribonacci system *where* $n_i = i$ *for all* $i \geq 1$ *and* $k_n = 1$ *for all* $n \geq 1$*; for this system, if* $y = 1,8392...$ *is the root bigger than* 1 *of the polynomial* $X^3 - X^2 - X - 1$*, then* $\mathcal{B} = 2y^2 + \frac{4y}{y^2+1} = 8,4445...$

Proof

We use the estimates from [11]; if n is an n_i, the smallest of the three lengths is $|J_n|$, and thus either $|H_{n_i}| \leq |G_{n_i}| \leq |H_{n_i}| + |J_{n_i}|$ or $|G_{n_i}| \leq |H_{n_i}| \leq |G_{n_i}| + |J_{n_i}|$, thus $|S_{n_i}| \leq 2|G_{n_i}|$ and $|S_{n_i}| \leq 2|H_{n_i}|$; thus the quantities $\frac{2|H_{n_i}|}{S_{n_i}}$ and $\frac{2|G_{n_i}|}{S_{n_i}}$ are between 1 and 2, while $\frac{2|J_{n_i}|}{S_{n_i}}$ is between 0 and 2.

Thus the finiteness of $\mathcal{B}$ depends on the two ratios $\frac{\theta_{n_i}}{\theta_{n_{i+1}+1}}$and $\frac{\theta_{n_i+1}}{\theta_{n_{i+1}+1}}$, and this gives the assertion on the highest value.

For the Tribonacci system, we get $|H_{n+1}| = |H_n| + |H_{n-1}| + |H_{n-2}|$, $|G_{n+1}| = |H_n| + |H_{n-1}|$, $|J_{n+1}| = |H_n|$, $\theta_n = \theta_{n+1} + \theta_{n+2} + \theta_{n+3}$, $\gamma_n = \theta_{n+1}$, $\iota_n = \theta_{n+2}$, and when n is large $|H_{n+1}|$ is close to $y|H_n|$ while θ_n is close to $y\theta_{n+1}$, which gives the formula.

Let us now sketch the proof that in all other cases $\mathcal{B} \geq \frac{181}{21}$. We suppose first that $k_m \geq 3$ for infinitely many m; we take such an m; let $r = n_i$ be the largest $n_i < m$, $s = n_{i+1}$.

- if $m = r + 1 < s$: then $\theta_{r+1} \geq \theta_{s+1}$, and $\theta_r \geq 3\theta_{r+1} + \theta_{r+2} + 3\theta_{s+1} \geq 3(\theta_{r+2} + \theta_{r+3}) + \theta_{r+2} + 3\theta_{s+1} \geq 10\theta_{s+1}$; putting all our estimates together, we get $\mathcal{B} \geq 11$,
- if $r + 1 < m \leq s$: then $\theta_{r+1} \geq \theta_{m-1} \geq 3\theta_m \geq 3\theta_{s+1}$, and $\theta_r \geq \theta_{r+1} + \theta_{r+2} + \theta_{s+1} \geq \theta_{m-1} + \theta_m + \theta_{s+1} \geq 4\theta_m + 2\theta_{s+1} \geq 6\theta_{s+1}$: we get $\mathcal{B} \geq 9$,

- if $m = r + 1 = s$, let $q = n_{i-1}$ and suppose $q < m - 2$; then $\theta_{q+1} \geq \theta_{m-1} + \theta_m \geq 4\theta_m$, and $\theta_q \geq \theta_{m-2} \geq 2(\theta_{m-1} + \theta_m) \geq 8\theta_m$, we get $\mathcal{B} \geq 12$,
- if $m = r + 1 = s$ and $k_m \geq 4$; then $\theta_{r+1} \geq \theta_{s+1}$, and $\theta_r \geq 4\theta_s + 4\theta_{s+1}$, we get $\mathcal{B} \geq 9$.

Hence either $\mathcal{B} \geq 9$, or k_n takes only the values $1, 2, 3$, with 3 possible only if $r = m - 1$ and $q = m - 2$ in the notations above. In this last case $\theta_{m-1} \geq 3(\theta_m + \theta_{m+1})$, and $\theta_{m-2} \geq \theta_{m-1} + \theta_m + \theta_{m+1} \geq 4(\theta_m + \theta_{m+1})$. If $k_{m+1} = 3$, we have $\theta_{m-1} \geq 9\theta_{m+1}$ and, computing it from $m - 1$, we get $\mathcal{B} \geq 9$. If $k_{m+1} \leq 2$, we have $\theta_m \leq 2\theta_{m+1} + 2\theta_{m+2} + \theta_{m+3}$ while $\theta_{m+2} \leq \theta_{m+1} - \theta_{m+3}$, hence $\theta_m \leq 4\theta_{m+1}$, and, computing it from $m-2$, we get $\mathcal{B} \geq \frac{35}{4} = 8,75$.

Thus we can suppose now that the k_n take only the values 1 and 2; then we can improve the estimates on the ratios of lengths, with $\frac{2|J_{n_i}|}{S_{n_i}} \geq \frac{1}{3}$, and, as the estimates from [11] imply that either $|S_{n_i}| \leq 2|G_{n_i}| - |J_{n_i}|$ or $|S_{n_i}| \leq 2|H_{n_i}| - |J_{n_i}|$, either $\frac{2|H_{n_i}|}{S_{n_i}}$ or $\frac{2|G_{n_i}|}{S_{n_i}}$ are greater than $\frac{8}{7}$; the lower bounds on the ratios $\frac{\theta_{m+1}}{\theta_m}$ are also improved. Thus we can prove that if for infinitely many m either $k_m = k_{m+1} = 2$, or $k_m = 2$ when m is not an n_i, or m is not an n_i and $m + 1$ is not an n_i, $\mathcal{B} \geq \frac{181}{21}$

Similarly, further improvement of the estimates, and extensive computations, allow us to eliminate all the $k_m = 2$ and all the m which are not an n_i. □

Theorem 4.4. *The smallest element in the lower BL spectrum of the family of Arnoux-Rauzy systems is* 2, *and the largest is* $+\infty$. *Every integer greater or equal to* 2 *is in the lower BL spectrum, and is an accumulation point, as is* $+\infty$.

Proof

The smallest element is at least 2 by Lemma 1.2. We take $n_i = 2i$, with $k_{2n+1} = 1$ and a sequence k_{2n} growing to $+\infty$. Then we have $|J_{2n}| < |G_{2n}| < |H_{2n}|$. $|H_{2n+2}| = k_{2n+2}(|H_{2n}| + |G_{2n}|) + |H_{2n}|$, $|G_{2n+2}| = k_{2n+2}(|H_{2n}| + |G_{2n}|) + |H_{2n}| + |J_{2n}|$, $|J_{2n+2}| = |H_{2n}| + |G_{2n}|$. When n tends to $+\infty$, $\frac{|G_{2n}|}{|H_{2n}|} \to 1$ and $\frac{|J_{2n}|}{|H_{2n}|} \to 0$, thus $\frac{|J_{2n}|}{|S_{2n+2}|} \to 0$ and both $\frac{|H_{2n}|}{|S_{2n+2}|}$ and $\frac{|G_{2n}|}{|S_{2n+2}|}$ are equivalent to $\frac{1}{4k_{2n+2}}$.

As for the ratios of frequencies, $\theta_{2n+1} = k_{2n+2}(\theta_{2n+2} + \theta_{2n+3}) + \theta_{2n+4} = k_{2n+2}(2\theta_{2n+3} + \theta_{2n+4} + \theta_{2n+5}) + \theta_{2n+4} \leq k_{2n+2}(2\theta_{2n+3} + 2\theta_{2n+4}) + \theta_{2n+4}$, and $\theta_{2n} = \theta_{2n+1} + \theta_{2n+2} + \theta_{2n+3} \leq (k_{2n+2} + 1)(2\theta_{2n+3} + 2\theta_{2n+4}) + \theta_{2n+4}$. As $\theta_{2n+4} \leq \frac{\theta_{2n+3}}{k_{2n+4}}$, both $\frac{\theta_{2n+1}}{\theta_{2n+3}}$ and $\frac{\theta_{2n}}{\theta_{2n+3}}$ are smaller than $2k_{2n+2}(1 + \epsilon_n)$, where $\epsilon_n \to 0$.

Putting everything together, we get that $\mathcal{B}'$ is at most 2, and thus equal to 2.

We take now $n_i = i$, with a sequence k_n tending to $+\infty$. Then for all n we have $|J_n| < |G_n| < |H_n|$, $S_{n+1} = (2k_{n+1} - 1)|H_n| + |G_n| + |J_n| < (2k_{n+1} + 1)|H_n|$, $\theta_n \geq k_{n+1}\theta_{n+1} \geq k_{n+1}k_{n+2}\theta_{n+2}$ and thus $\mathcal{B}'$ is at least the lower limit of $\frac{2k_{n+1}k_{n+2}}{2k_{n+1}+1}$, which gives $+\infty$.

If we replace $k_{2n+1} = 1$ by $k_{2n+1} = m$ in the first construction of the proof above, we get $\mathcal{B}' = m + 1$, for any integer $m \geq 1$.

If we replace the variable k_n or k_{2n} in the constructions above by a constant k and let k tend to infinity, we get sequences in the lower spectrum tending to $+\infty$ or to $m + 1$, for any integer $m \geq 1$. □

Thus we have a new, and very simple, family of counter-examples to the necessity of Boshernitzan's criterion.

Of course, there are other values in the lower spectrum than those in Theorem 4.4, and we conjecture that *the lower BL spectrum of the family of Arnoux-Rauzy systems is the interval* $[2,+\infty]$.

We look at some individual values of $\mathcal{B}$ and $\mathcal{B}'$, for three examples which are codings of rotations of the 2-torus [1] [2].

For the Tribonacci system $\mathcal{B}' = 2y + \frac{4}{y^2+1} = 4,5911....$ In [13], various constants are computed for the Tribonacci system, including one covering number, though not by intervals (see Section 2 above), and a measure of the quality of simultaneous approximation of $(\frac{1}{y}, \frac{1}{y^2})$ by rationals; none of these is equal to $\mathcal{B}$ or $\mathcal{B}'$.

For $n_i = i$ and $k_n = 2$, we get $|H_{n+1}| = 2|H_n| + 2|H_{n-1}| + |H_{n-2}|$, $|G_{n+1}| = |2H_n| + |H_{n-1}|$, $|J_{n+1}| = |H_n|$, $\theta_n = 2\theta_{n+1} + 2\theta_{n+2} + \theta_{n+3}$, and when n is large H_{n+1} is close to $y_0 H_n$ while θ_n is close to $y_0\theta_{n+1}$, if $y_0 = 2,83...$ is the root bigger than 1 of the polynomial $X^3 - 2X^2 - 2X - 1$. We get $\mathcal{B} = 2y_0^2 + \frac{4y_0}{y_0^2+y_0+1} = 16,96...$, $\mathcal{B}' = 2y_0 + \frac{4}{y_0^2+y_0+1} = 5,99...$

For $n_i = 2i$ and $k_n = 1$ for all n, we get $|H_{2n}| = |H_{2n-1}| + |H_{2n-2}|$, $|H_{2n-1}| = |H_{2n-2}| + |H_{2n-3}| + |H_{2n-4}| + |H_{2n-5}|$, $|G_{2n}| = |H_{2n-1}| + |H_{2n-2}| + |H_{2n-3}|$, $|J_{2n}| = |H_{2n-1}|$, $\theta_{2n} = \theta_{2n+1} + \theta_{2n+2} + \theta_{2n+3}$, $\theta_{2n+1} = \theta_{2n+2} + \theta_{2n+3} + \theta_{2n+5}$. Let $y_1 = 1,4516...$ be the root bigger than one of the polynomial $2X^3 - X^2 - 2X - 1$, $y_2 = \frac{1}{y_1-1} = 2,2143...$, $y_4 = \frac{1}{y_1} + 1 = 1,689...$ and $y_3 = 1,903...$ be the number bigger than one satisfying $y_3^2 y_4 = y_3 y_4 + y_3 + 1$. When n is large, $|H_{2n}|$ is close to $y_1|H_{2n-1}|$, $|H_{2n-1}|$ is close to $y_2|H_{2n-2}|$, θ_{2n} is close to $y_3\theta_{2n+1}$, θ_{2n+1} is close to $y_4\theta_{2n+2}$. Then $\mathcal{B} = \frac{2y_1^2 y_2}{y_1^2 y_2 + y_1 + 1}(y_3^2 y_4 + \frac{y_1^2 y_2 y_3}{y_1 y_2 + y_1 + 1} + \frac{1}{y_1}) = 11,61...$, which is a candidate for the second smallest value in the upper spectrum, and $\mathcal{B}' = \frac{\mathcal{B}}{y_1 y_2} = 3,61...$

The Arnoux-Rauzy systems raise questions about rotations of the 2-torus, and we may ask what could be the BL spectra for that family of systems, but the problem is that at this time we do not know any coding which may be called natural. In Section 3 of [6], there is a discussion about what should be the properties of such a coding; in view of the present paper, by analogy with the rotations of the 1-torus, it becomes reasonable to add to these properties the boundedness of the lower BL spectrum these natural codings would define for the family of rotations of the 2-torus.

Now, if we code a rotation of the 2-torus with the Cartesian product of two partitions of the 1-torus, then the complexity is quadratic and all $\mathcal{B}$ and $\mathcal{B}'$ are infinite by Lemma 1.2, which gives another trivial counter-example to the necessity of Boshernitzan's criterion, but this coding has never been considered as natural. The Arnoux-Rauzy systems were devised to provide codings with sub-linear complexity for rotations of the 2-torus, but this was succesful only in a limited number of cases. Still, if we consider these cases, the tentative lower BL spectrum of the family of rotations of the 2-torus seems to be quite different from the lower BL spectrum of rotations of the 1-torus: if we take an Arnoux-Rauzy system with $n_i = i$ and constant $k_n = k$, it is a coding of a rotation of the 2-torus by [2], and these give arbitrarily high values for $\mathcal{B}'$; if $n_i = i$ and k_n grows slowly (for example $k_n \leq \frac{1}{15}n$), we get an infinite $\mathcal{B}'$ while the Arnoux-Rauzy system is shown in [11] to have two continuous eigenfunctions, and is still conjectured to be a coding of a rotation of the 2-torus. Thus it seems that even those Arnoux-Rauzy which do code rotations of the 2-torus fail to satisfy our new condition for being natural codings, and thus we need new ideas to find natural codings of these rotations.

5. Questions

N. Pytheas Fogg (private communication) asked what should be the BL spectra of the family of all uniquely ergodic symbolic systems. The upper BL spectrum contains the union of the Lagrange spectrum and $+\infty$, and it is quite possible that it is just that. The lower BL spectrum contains the lower BL spectrum of rotations, and the interval $[2, +\infty]$, but also at least the point $\frac{3}{2}$, which is the value of $\mathcal{B}'$ for the so-called *period-doubling* symbolic system, whose language is generated by the fixed point of the *substitution* $a \to ab$, $b \to aa$ (N. Pytheas Fogg, unpublished).

N. Pytheas Fogg has also started to investigate the spectrum of the joint values of $(\mathcal{B}, \mathcal{B}')$ for rotations, see https://www.lirmm.fr/ monteil/hebergement/pytheas-fogg/BL_spectrum.pdf.

References

[1] P. ARNOUX: Un exemple de semi-conjugaison entre un échange d'intervalles et une translation sur le tore, *Bull. Soc. Math. France* 116 (1988), p. 489–500.

[2] P. ARNOUX, Sh. ITO: Pisot substitutions and Rauzy fractals. Journées Montoises d'Informatique Théorique (Marne-la-Vallée, 2000), *Bull. Belg. Math. Soc. Simon Stevin* 8 (2001), p. 181–207.

[3] P. ARNOUX, G. RAUZY: Représentation géométrique de suites de complexité $2n+1$, *Bull. Soc. Math. France* 119 (1991), p. 199–215.

[4] V. BERTHÉ: Fréquences des facteurs des suites sturmiennes. (French) [Frequencies of the factors of Sturmian sequences], *Theoret. Comput. Sci.* 165 (1996), p. 295–309.

[5] V. BERTHÉ, N. CHEKHOVA, S. FERENCZI: Covering numbers: arithmetics and dynamics for rotations and interval exchanges, *J. Anal. Math.* 79 (1999), p. 1–31.

[6] V. BERTHÉ, S. FERENCZI, L. ZAMBONI: Interactions between dynamics, arithmetics and combinatorics: the good, the bad, and the ugly, *Algebraic and topological dynamics, Contemp. Math.* 385 (2005), p. 333–364, Amer. Math. Soc., Providence, RI.

[7] M. BOSHERNITZAN: A unique ergodicity of minimal symbolic flows with linear block growth, *J. Analyse Math.* 44 (1984/85), p. 77–96.

[8] M. BOSHERNITZAN: A condition for minimal interval exchange maps to be uniquely ergodic, *Duke Math. J.* 52 (1985), p. 723–752.

[9] M. BOSHERNITEZAN: A condition for unique ergodicity of minimal symbolic flows, *Ergodic Theory Dynam. Systems* 12 (1992), p. 425–428.

[10] J. CASSAIGNE: Limit values of the recurrence quotient of Sturmian sequences. WORDS (Rouen, 1997). *Theoret. Comput. Sci.* 218 (1999), p. 3–12.

[11] J. CASSAIGNE, S. FERENCZI, A. MESSAOUDI: Weak mixing and eigenvalues for Arnoux-Rauzy sequences. *Ann. Inst. Fourier (Grenoble)* 58 (2008), p. 1983–2005.

[12] N. CHEKHOVA: Covering numbers of rotations, *Theoret. Comput. Sci.* 230 (2000), p. 97–116.

[13] N. CHEKHOVA, P. HUBERT, A. MESSAOUDI: Propriétés combinatoires, ergodiques et arithmétiques de la substitution de Tribonacci. (French) [Combinatorial, ergodic and arithmetical properties of the Tribonacci substitution], *J. Théor. Nombres Bordeaux* 13 (2001), p. 371–394.

[14] T. CUSICK, M. FLAHIVE: The Markoff and Lagrange spectra.*Mathematical Surveys and Monographs*, 30. American Mathematical Society, Providence, RI, 1989. x+97 pp.

[15] S. FERENCZI, C. HOLTON, L. ZAMBONI: The structure of three–interval exchange transformations I: an arithmetic study, *Ann. Inst. Fourier (Grenoble)* 51 (2001), p. 861–901.

[16] S. FERENCZI, C. HOLTON, L. ZAMBONI: The structure of three–interval exchange transformations II: a combinatorial description of the trajectories, *J. Anal. Math.* 89 (2003), p. 239–276.

[17] S. FERENCZI, C. HOLTON, L. ZAMBONI: The structure of three-interval exchange transformations III: ergodic and spectral properties, *J. Analyse Math.* 93 (2004), p. 103–138.

[18] S. FERENCZI, T. MONTEIL: Infinite words with uniform frequencies, and invariant measures, in *Combinatorics, automata and number theory*, Encyclopedia of Mathematics and its Applications (No. 135), Cambridge University Press (2010), ISBN: 9780521515979, p. 373–409.
[19] S. FERENCZI, L.F.C. da ROCHA: A self-dual induction for three-interval exchange transformations, *Dyn. Syst.* 24 (2009), p. 393–412.
[20] S. FERENCZI, L. ZAMBONI, Structure of k-interval exchange transformations: induction, trajectories, and distance theorems, *J. Anal. Math.* 112 (2010), p. 289–328.
[21] S. FERENCZI, L. ZAMBONI: Languages of k-interval exchange transformations, *Bulletin London Math. Soc.* 40 (2008), p. 705–714.
[22] C. GRILLENBERGER: Constructions of strictly ergodic systems. I. Given entropy. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 25 (1972/73), p. 323–334.
[23] M. HALL Jr. On the sum and product of continued fractions. *Ann. of Math. (2)* 48, (1947), p. 966–993.
[24] N. PYTHEAS FOGG: Substitutions in dynamics, arithmetics and combinatorics. Edited by V. Berth, S. Ferenczi, C. Mauduit and A. Siegel. Lecture Notes in Mathematics, 1794. Springer-Verlag, Berlin, 2002. xviii+402 pp. ISBN: 3-540-44141-7
[25] G. RAUZY: Nombres algébriques et substitutions, *Bull. Soc. Math. France* 110 (1982), p. 147–178.
[26] W. A. VEECH: Boshernitzan's criterion for unique ergodicity of an interval exchange transformation, *Ergodic Theory Dynam. Systems* 7 (1987), p. 149–153.
[27] W. A. VEECH: Measures supported on the set of uniquely ergodic directions of an arbitrary holomorphic 1-form. *Ergodic Theory Dynam. Systems* 19 (1999), p. 1093–1109.

Institut de Mathématiques de Luminy, CNRS - UMR 6206, Case 907, 163 av. de Luminy, F13288 Marseille Cedex 9 (France), and Fédération de Recherche des Unités de Mathématiques de Marseille, CNRS - FR 2291

E-mail address: ferenczi@iml.univ-mrs.fr